



TAMPEREEN TEKNILLINEN YLIOPISTO

ANTTI KOIVISTO

KONTEKSTITIETOON PERUSTUVA AUTOMAATTINEN TAGIEN
LUONTI MOBIILILAITTEILLA

Diplomityö

Tarkastaja: professori Jari Multisilta
Tarkastaja ja aihe hyväksytty
Tieto- ja sähkötekniikan
tiedekuntaneuvoston
kokouksessa 13. tammikuuta 2010

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO

Tietotekniikan koulutusohjelma

Koivisto, Antti: Kontekstitietoon perustuva automaattinen tagien luonti mobiililaitteilla

Diplomityö, 67 sivua, 3 liitesivua

Tammikuu 2010

Pääaine: Ohjelmistotuotanto

Tarkastaja: professori Jari Multisilta

Avainsanat: Tagi, kontekstitieto, tagikirjasto, mobiililaite, automaattinen tagi, mobiili video

Videoiden jakaminen on saanut suurta suosiota sosiaalisen median palveluissa. Esimerkiksi YouTubeen, joka on yksi suosituimpia videopalveluita maailmassa, ladataan satojatuhansia videoita päivittäin. Monet ladatuista videoista on kuvattu käyttäen mobiililaitetta, kuten matkapuhelinta. Tämä johtuu osaksi siitä, että matkapuhelimilla pystytään nykyään kuvaamaan melko hyvänlaatuisia videoita ja matkapuhelin on lähes aina käyttäjän mukana.

Matkapuhelimella kuvatun videon lähettäminen videopalveluun on kuitenkin vielä melko työlästä, koska video on useimmissa palveluissa aluksi siirrettävä tietokoneelle, josta se voidaan siirtää videopalveluun. Videoiden suuresta määrästä johtuen videoiden löytäminen videopalvelusta on hankalaa. Tampereen teknillinen yliopisto on kehittänyt MoViE –palvelun (Mobile Video Experience) juuri tällaisia tilanteita varten. Palvelu sisältää sivuston, joka sisältää ladatut videot, sekä client-ohjelmiston. Client-ohjelmisto avustaa videon lataamisessa palveluun, jos video halutaan lähettää videopalveluun heti kuvaamisen jälkeen.

Koska videoiden määrä palveluissa kasvaa jatkuvasti, niiden löytämiseksi on syötettävä avainsanoja, jotka avustavat videoiden hakua. Avainsanojen eli tagien syöttäminen voi olla joillekin käyttäjille hankalaa sekä aikaa vievää varsinkin matkapuhelimella, johtuen matkapuhelinten pienistä ja toisistaan eroavista näppäimistöistä. Tämän vuoksi moniin palveluihin on pyritty toteuttamaan tagikirjasto, joka ehdottaa käyttäjälle sopivia tageja.

MoViE -palveluun rakennettu tagikirjasto perustuu matkapuhelimesta saatavaan kontekstitietoon. Kontekstitiedon avulla voidaan tehdä tietokantakyselyjä ja hakea tietoa Internetistä, mikä taas mahdollistaa mahdollisimman tarkkojen tagiehdotusten antamisen käyttäjälle. Client-ohjelma tulostaa saadut tagiehdotukset matkapuhelimen näytölle, josta sopivimmat tagit on helppo valita. Tagikirjasto on rakennettu siten, että se soveltuu sekä yksittäisiin videoihin sekä tapauksiin, joissa palveluun lähetetään monta videota lyhyen ajan sisällä.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY

Master's Degree Programme in Information Technology

Koivisto, Antti: Automatic tag generation with mobile devices based on context information

Master of Science Thesis, 67 pages, 3 Appendix pages

January 2010

Major: Software engineering

Examiner: Professor Jari Multisilta

Keywords: Tag, context information, taglibrary, mobile device, automatic tag, mobile video

Video sharing has gained a major role among social media applications. For example hundreds of thousands of videos are daily uploaded to YouTube, which is one of the most popular video sharing services in the world. Many of the uploaded videos are filmed by using a mobile device such as cellphone. This is partly because users are able to film videos of quite good quality using cellphones and because users carry cellphones with them almost everywhere.

It is still quite tedious to send mobile video to video sharing service. In most services, video must be transferred first to the computer and after that it can be transmitted to video sharing service. The large number of videos makes it difficult to find specific video from a video sharing service. The Tampere University of Technology has developed MoViE –service (Mobile Video Experience) for situations like this. Service includes a site, which consists of uploaded videos and a client program. The client program collaborates with a video uploading to a service, if a user prefers to send a video to the video sharing service right after capturing a video.

As the number of videos in the service continues to grow, users must enter keywords to assist a video retrieval. Keyword alias tags feeding may be difficult and time-consuming for some users, especially with cellphones, due to the small size of the keypad and the difference between various keypads. Because of this, many services have tried to develop a tag library. The tag library proposes suitable tags for the user.

The tag library built in the MoViE –service is based on cellphones' contextual information. With context information it is possible to make database queries and search

information from Internet, which on the other hand makes it possible to give accurate tag suggestions to the users. The client program prints tag suggestions to the cellphones display where the user can easily choose the most suitable tags. The tag library is built in a way that it supports both single videos and situations where the user sends many videos in a short time.

ALKUSANAT

Tämä diplomityö on tehty Tampereen teknillisen yliopiston (TTY) Porin yksikölle AMC:lle (Advanced Multimedia Center) osana MoViE (Mobile Social Media: video applications for entertainment and learning) – projektia. Työn tekeminen on ollut erittäin mielenkiintoista ja välillä hieman haastavaakin. Työ oli kuitenkin erittäin antoisa ja mielenkiintoinen kokemus.

Haluan kiittää erityisesti työn tarkastajaa professori Jari Multisiltaa mahdollisuudesta tehdä tämä diplomityö. Hänen kannustava asenteensa sekä hyvät neuvot ovat auttaneet työn tekemisessä.

Ohjaaja Arttu Perttulaa haluan kiittää opastuksesta, kannustuksesta ja asiantuntevista neuvoista työn eri vaiheissa.

Lisäksi haluan kiittää tutkija Marko Suomista sekä dosentti Harri Ketamoa heidän asiantuntevista näkökulmistaan ja korvaamattomasta avusta työn kehittymisen kannalta. Haluaisin kiittää myös työkavereita tutkija Pauliina Tuomea sekä vanhempaa tutkijaa Kristian Kiiliä avusta työn loppuun saattamisessa sekä koko työporukkaa mukavista työpäivistä ja motivoinnista. Myös Pori Jazzin henkilökunta on antanut mielenkiintoisia näkökulmia työn kehittämisen kannalta.

Avopuolisoani Saria haluan kiittää korvaamattomasta tuesta niin työn edistymisen kannalta, kuin myös siviilielämässä.

Äitiäni Tuulaa, isääni Markkua, siskoani Paulaa sekä ystäviäni haluan kiittää tuesta ja kannustuksesta tätä työtä tehdessä ja koko opiskeluaikanani.



Euroopan unioni
Euroopan aluekehitysrahasto

Porissa 17.01.2010

Antti Koivisto
Mattilantie 1
28220 Pori

SISÄLLYS

1. Johdanto	1
2. Tieto sähköisessä mediassa	4
2.1. Konteksti- ja metatieto	5
2.2. Tagi	7
2.2.1. Manuaalinen tagi	9
2.2.2. Automaattinen tagi	10
2.2.3. Tagikirjasto	11
2.3. Geotagi	12
2.3.1. GPS -paikannus	12
2.3.2. Solupaikannus	13
2.3.3. Verkkoavusteinen paikannus	14
3. Tiedon liittäminen sähköiseen- ja mobiili mediaan	16
3.1. Tiedon liittäminen kuvaan	16
3.2. Kontekstitiedon liittäminen videoon	18
4. Tagi ja kontekstitieto mobiililaitteessa	21
4.1. Sovelluksia	21
4.1.1. ZoneTag	22
4.2. Case: MoViE –projekti Pori Jazzeilla	25
4.2.1. Case I: Kilpailu Pori Jazzeilla	25
4.2.2. Case II: MoViE –client testaus	27
5. Taggaus MoViE -projektissa	30
5.1. Taustatutkimukset	31
5.1.1. Tagikirjaston hyödyllisyys	31
5.2. Tiedon kerääminen	33
5.2.1. Palvelin ja tietokannan tekninen toteutus	34
5.2.2. Mobiililaite	36
5.3. Tiedon liittäminen ja siirtäminen	37
5.3.1. Tagipalvelu palvelimella	38
5.3.2. Tagipalvelu mobiililaitteessa	38
5.4. Tagipalvelun tekninen toteutus	38
5.4.1. Tekninen toteutus palvelin ja client	40
5.4.2. Tulevaa MoViE –palveluun	52
5.5. Sovellusten yhteensopivuus	53
6. Nykytilanne ja tulevaisuus	54
7. Yhteenvedo	56

Lähteet	58
Liitteet	
Liite 1	65

LYHENTEET

3G	Third generation, yleinen lyhenne ns. "kolmannen sukupolven" matkapuhelinteknologioille.
3GP	Formaatti multimedian pakkaukseen.
A-GPS	Assisted GPS tarkoittaa avustettua satelliittipaikannusta.
EDGE	Matkapuhelinten pakettipohjaiseen tiedonsiirtoon suunniteltu tekniikka, perustuu GPRS-tekniikkaan.
GNSS	Global Navigation Satellite System, tarkoittaa eri satelliittipaikannusjärjestelmiä.
GPRS	General Packet Radio Service, GSM-verkossa toimiva pakettikytkentäinen tiedonsiirtopalvelu.
GPS	Global Positioning System, satelliittipaikannusjärjestelmä.
HD	High definition, teräväpiirtotekniikka.
ID	Identification, tunnistus.
Mobile Python	Python ohjelmointikieli matkapuhelimille.
MoViE	Mobile Video Experience, projekti johon diplomityö tehtiin.
MySQL	SQL-tietokannan hallintajärjestelmä, Structured Query Language (SQL).
Navstar GPS	All-weather navigation system, paikannusjärjestelmä.
PDA	Personal Digital Assistant, kämmentietokone.

PHP	Hypertext Preprocessor, ohjelmointikieli.
PyS60	Katso Mobile Python.
RSS	Really Simple Syndication, verkkosyötemuoto, jota käytetään usein päivittyvän digitaalisen sisällön julkaisemiseen.
S60-sarja	Nokian kehittämä matkapuhelinten käyttöliittymä ja sovellusalusta.
T9	Text on 9 keys, Tegic Communications -yrityksen kehittämä matkapuhelimissa käytetty ennakoiva tekstinsyöttöjärjestelmä.
TA (Timing Advance)	Ajastusennakko.
WLAN	Wireless Local Area Network, langaton lähiverkko.
WWW	World Wide Web, Internet-verkossa toimiva hajautettu hypertekstijärjestelmä.
XML	eXtensible Markup Language, metakieli jonka avulla kuvataan tietoa tiedosta.

SANANSELITYKSET

Banneri	Www-sivuilla olevia pieniä mainoksia.
Blogi	Verkkoloki (engl. blog, weblog) on julkinen säännöllisesti päivitetty verkkosivusto.
Case	Kuvaus todellisesta tilanteesta tai ongelmasta, joita joudutaan ratkaisemaan.
Client	Palvelua käyttävä sovellus.
Duplikaatti	Jäljennös, kaksoiskappale.
Facebook	Internetissä toimiva yhteisöpalvelu.
Flash	Adobe Systemsin tuottama kehitysympäristö, jonka avulla voidaan luoda multimediaesityksiä.
Flickr	Webin ensimmäinen ilmainen valokuvien säilytys- ja jakopalvelu.
Flyer	Lentolehtinen.
Folksonomia	Yhteisöllisesti tuotettu, avoin luokittelujärjestelmä.
Generoida	Tuottaa, muodostaa, kehittää.
Geokoodaus	Yleisnimi analyyseille, joiden avulla osoite-, postinumero- tai muun tunnistetiedon omaaville tietueille luodaan spatiaalisen tiedon sisältävien tietueiden avulla maantieteellinen sijainti.
Geotagi	Maantieteelliseen sijaintiin liitetty tagi.

Google	Google-verkosto koostuu suuresta määrästä sivustoja ja muita tuotteita.
GSM solu(verkko)	Tukiaseman peittämä maantieteellinen alue.
Indeksointi	Aputietorakenteita, jotka nopeuttavat tiedonsaantia.
JavaScript	Web-ympäristössä käytettävä komentosarjakieli.
Kontekstitieto	Tapahtumaan tai asiaan liittyvää lisäinformaatiota.
Konvertointi	Muuntaminen, esimerkiksi tiedoston muuntaminen eri tyyppiseksi.
Live stream	Verkossa tapahtuvaa reaaliaikaista jatkuvaa tiedonsiirtoa.
Manuaalinen tagi	Käyttäjän kirjoittama tagi.
Metatieto	Kuvailevaa ja määrittävää tietoa jostakin tietovarannosta tai sisältöyksiköstä.
Mikroblogi	Blogi, jonka merkinnät ovat lyhyitä, tavallisesti korkeintaan 140 merkin mittaisia.
Mobiililaite	Kannettava laite, jolla voidaan siirtää tietoa langattomasti, esimerkiksi matkapuhelin tai PDA.
Ontologia	Luokitteluja, joita on käytetty erityisesti automaattisen tietojenkäsittelyn yhteydessä.
Palvelin	Tietokoneessa suoritettava palvelinohjelmisto sekä ohjelmistoa suorittava tietokone.
Prototyyppi	Alkuperäinen, ensimmäinen versio.
Python	Monipuolinen, tulkattava ohjelmointikieli.
Päätelaite	Tietoliikenneyhteyden päässä oleva laite, jonka avulla varsinainen kommunikointi tietoliikenneyhteyden yli tapahtuu.

Semanttinen web	Internetin WWW-palvelun laajennus, jonka dokumentit on suunniteltu myös koneita silmällä pitäen.
Serveri	Katso palvelin.
Sosiaalinen media	Prosessi, jossa yksilöt ja ryhmät rakentavat yhteisiä merkityksiä sisältöjen, yhteisöjen ja verkkoteknologioiden avulla.
Symbian	Pienitehoisille ja vähillä resursseilla varustetuille laitteille tarkoitettu käyttöjärjestelmä.
Tagi	Avainsanoja, jotka kuvaavat asian sisältöä.
Tagiavaruus	Avainsanojen joukko.
Tagikirjasto	Järjestetty tagiavaruus, käytetään tagien ehdottamiseen.
Tagipilvi	Graafinen versio tagiavaruudesta, jossa korostetaan useimmiten esiintyviä tageja.
Taksonomia	Yleisnimi tieteele, joka tutkii samojen asioiden käytännön puolta.
Taulu	Tietokannan sisältämä pienempi informaatio yksikkö.
Tausta-ajo	Taustalla, käyttäjältä huomaamattomasti, tapahtuva suoritus.
Tietokanta	Tietovarasto, kokoelma tietoja, joilla on yhteys toisiinsa.
Tukiasema	Langattomassa tietoliikenteessä päätelaitteen yhdistäminen radioteitse kiinteään verkkoon.
Twitter	Yhteisö- ja mikroblogipalvelu, jonka käyttäjät pystyvät lähettämään ja lukemaan toistensa päivityksiä Internetissä.
Uniikki	Ainoa lajiaan, ainutlaatuinen.
Unix	Laitteistoriippumaton käyttöjärjestelmä.

Videopalvelu	Palvelu, joka tarjoaa käyttäjille videoita, vertaa YouTube.
Yahoo	Tarjoaa suuren määrän eri palveluja käyttäjälle.
ZoneTag	ZoneTag on ohjelma, joka mahdollistaa kuvien tallennuksen, huomautuksien lisäämisen, varastoinnin ja lataamisen suoraan matkapuhelimella.
Älypuhelin	matkapuhelin, jossa on perinteisten mobiilipuhelintoimintojen lisäksi kämmentietokonetta muistuttavia ominaisuuksia.

1. JOHDANTO

Älypuhelimet ovat yleistyneet viime vuosina. Tämä on mahdollistanut laitekeskeisten sovellusten kehittymisen, paremman suorituskykynsä ansiosta. Myös matkapuhelimien parantuneet näytöt ovat oleellinen osa tätä kehitystä, sekä päätelaiteympäristön avautuminen. [1] Puhelimeissa on myös yleisesti mahdollisuus verkkotyöskentelyyn ja niillä on mahdollista tallentaa mediaa. Kun yhdistetään media ja kyberinfrastruktura, avautuu monia kiinnostavia mahdollisuuksia videon avulla kerrontaan ja videolla keskusteluun olennaisena osana hajautettuja luovia ympäristöjä. Tämä työ toteutetaan osana MoViE (Mobile Social Media: video applications for entertainment and learning) – projektia, joka pyrkii kehittämään ja testaamaan uutta sosiaalista mobiilivideopalvelua taiteen, viihteen ja oppimisen alueilla. Projektissa työskentelevät tutkijat ovat Suomesta. Projekti on kaksivuotinen ja sen päärahoittajana toimii Tekes sekä muut rahoittajat ovat Satakunnan kirjateollisuus Oy, Porin Seudun Matkailu Oy, Floobs Oy, Pori Jazz 66 ry ja Qwertomec Oy. Projektin vastuullisena johtajana toimii professori Jari Multisilta Tampereen teknillisestä yliopistosta. Professori Jaakko Suominen ja dekaani Marjo Mäenpää vastaavat projektin toteutuksesta Turun yliopiston ja Taideteollisen korkeakoulun osalta.

Tässä työssä käsitellään kontekstietoon perustuvaa taggausta mobiililaitteilla, mikä tarkoittaa automaattisesti saatavan tiedon liittämistä tallenteeseen matkapuhelimella. Työssä pyritään etsimään vastausta ongelmaan, miten tageja voidaan antaa tallenteeseen ja mitkä voisivat olla sopivia tageja juuri kyseiseen tallenteeseen. Kontekstietona puhelimeen voidaan asettaa muun muassa käyttäjän tiedot, aika, paikka ja sää. Näiden perusteella saadaan paljon tietoja, joiden mukaan voidaan ehdottaa erilaisia avainsanoja eli tageja käyttäjälle. Yksi tämän työn tärkeimpiä asioita on luoda automaattisten tagisanojen ehdotus käyttäjälle. Tagien lisääminen pyritään ratkaisemaan eri tavalla kuin monissa muissa tunnetuissa tageja käyttävissä sovelluksissa. Tagit pyritään liittämään tallenteeseen mahdollisimman automaattisesti, esimerkiksi ehdottaen todennäköisimpiä tageja, joita käyttäjä voisi kyseisessä paikassa kyseiseen aikaan käyttää.

Sosiaalinen verkostoituminen ja Internet ovat tulleet suureksi osaksi kommunikointia ja verkostoitumista sekä liiketoiminnassa että viihdykkeenä. Jo nyt on olemassa monia suuria verkostoitumissivustoja. Näihin voidaan laskea mm. bloggaus sivusto Blogger sekä kuvien ja videoiden jakamissivustot YouTube, Facebook, Flickr sekä yritysmaailman verkostoitumispalvelu LinkedIn. Kyseisissä palveluissa kuvia ja videoita voidaan etsiä avainsanoilla eli tageilla. Kuvalla tai videolla voi olla useita tageja, mutta yleensä kuitenkin tageja määritellään nolasta kahteenkymmeneen.

Kuvilla voi olla myös geotageja, jotka määrittävät, missä kuva on otettu. Uusimmissa puhelimissa on automaattinen tuki geotagien käytölle. Tällainen sovellus liittää kuvaan lisätietoa, joka sisältää kuvan ottopaikan paikkakoordinaatit. Tuki toimii, jos se on sallittu ja gps tiedot on saatavilla.

Projektiin kuuluu edellä mainittujen sovelluksien omaksumisen konkreettinen analysointi. Lisäksi projektissa tutkitaan, kuinka käytännöllisiä ja kiinnostavia sovellukset jokapäiväisessä käytössä ovat. Osana tutkimusta selvitetään, miten teknologian ja muiden ratkaisuiden tulisi kehittyä, jotta sovellukset yleistyisivät valtavirran käyttöön. Näitä pyritään selvittämään kahdella alueella useilla tapaustutkimuksilla. Alueisiin kuuluvat viihdetapahtumat ja oppimistapahtumat. Tapaukset toteutetaan Suomessa (Pori). Viihdetapahtumana käytetään Pori Jazz –festivaalia, jossa ihmiset kokoavat omia kertomuksiaan mobiililaitteilla. Oppimistapahtumat koostuvat erilaisista mobiililuokkaretkistä Suomessa.

Tämän diplomityön tarkoituksena on rakentaa projektiin ohjelmiston osa, joka mahdollistaa kontekstitiedon lisäämisen videoon. Kontekstitietona pidetään kyseisessä projektissa mm. aikaa, paikkaa, säätä ja omia tageja, jotka käyttäjä lisää itse. Ohjelmiston yhtenä ominaisuutena tulee olemaan tagikirjasto. Tagikirjasto ehdottaa automaattisesti erilaisia avainsanoja, joita kuvaan voisi liittää. Ohjelmisto toteutetaan mobiililaitteessa Python- ja palvelimessa PHP-ohjelmointikoodilla. Ohjelmisto hyödyntää aikaisemmin syötettyjä tageja, joita se ehdottaa käyttäjille tilanteesta riippuen. Tähän syvennyttään vielä myöhemmin kappaleessa 5. Taggaus MoViE -projektissa.

Tämä työ toteutettiin osana sosiaalista MoViE (Mobile Video Experience) mobiilipalvelua, joka mahdollistaa käyttäjiä luomaan videokertomuksia matkapuhelimilla. MoViE on suunniteltu matkapuhelimille ja sen avulla voidaan lisätä tageja, tagiavaruuksia ja geotageja videoihin. MoViE –palveluun on myös rakennettu valvontamahdollisuus ei toivottuja tilanteita varten. Ajatus MoViE:ssä on yksinkertainen. Se mahdollistaa käyttäjien matkapuhelimella kuvaamien videoiden lähetyksen palveluun. Ensi näkemältä MoViE saattaa näyttää perinteiseltä videoidenjakosivustolta. Kun palvelua tarkastelee lähemmin voi huomata, että se on suunniteltu tutkimusalustaksi, jonka avulla on tarkoitus selvittää, kuinka ihmiset voivat luoda tarinoita sosiaalisen median mobiilipalvelulla. Sitä sovelletaan myös tutkimaan mobiilioppimista sosiaalisen median ja videon avulla. MoViE ei tue toistaiseksi reaaliaikaista videonjakoa, koska tällöin ei olla riippuvaisia Internet-yhteyden nopeudesta tai –kattavuudesta. [2; 3]

Monet käytössäolevat videoidenjakopalvelut on kehitetty tietokonekäyttöön. Käyttäjät voivat ladata palveluihin videoita matkapuhelimillaan, mutta muut toiminnot on tehtävä palveluun tarkoitettulla Internet-sivustolla. Tämä johtuu yleisimmin JavaScript:n käytöstä, koska suurin osa matkapuhelimista ei tarjoa täyttä tukea sille. MoViE –sivuston suunnittelussa on kuitenkin otettu huomioon matkapuhelimen selaimen rajoitukset. Ulkoasu on tehty mahdollisimman yksinkertaiseksi ja kevyeksi. Toiminta on toteutettu pääosin palvelimella ja käyttäjän antamia syötteitä on

yksinkertaistettu. Esimerkiksi videon taggaaminen onnistuu helposti valitsemalla sopivimmat tagit valikosta. Tyypillisen Flash-soittimen lisäksi videot on saatavilla myös 3gp –muodossa. Jotta latausajat olisivat siedettäviä, videon pituus on rajoitettu hyvällä kuvanlaadulla noin minuuttiin, riippuen videon koosta. Jakob Nielsenin verkkosuunnittelun klassikoksi muodostuneessa *Designing Web Usability* –teoksessa vuonna 2000 suhtauduttiin verkkovideoinnin tekemiseen vielä hyvin negatiivisesti [4]. Nielsenin ja Lorangerin (2006) tuoreempi arvio on kuitenkin jo myönteisempi. Siinä on todettu verkossa olevan videon ihannemitan olevan yleensä alle minuutin, muutoin katsojat eivät jaksakaan keskittyä videoon. Myös käytettävyys on kehittynyt muutamassa vuodessa huomattavasti. [5] Nielsenin ja Lorangerin toteamus videoiden pituudesta soveltuukin varsin mainiosti MoViE:n videoiden normaaliin mittaan. [2; 3]

MoViE sisältää tyypillisimmät videoidenjakosivustojen toiminnot. Käyttäjät voivat siirtää omia videoita palveluun, katsoa ja arvostella videoita, sekä vastata videoon omalla videollaan. Videolista voidaan lajitella päivämäärän mukaan, videon pisteiden mukaan tai sen mukaan, kuinka monta kertaa video on katsottu. Palvelussa voidaan myös etsiä videoita esimerkiksi käyttäjän, otsikon, kuvauksen, tagien tai geotagien mukaan. Käyttäjä voi myös lisätä tageja omiin videoihinsa jälkikäteen MoViE -sivustolla. Monissa tapauksissa on sopivaa järjestellä omat tagit tagiavaruuteen, joka on kokoelma käytettyjä avainsanoja tietyltä käyttäjältä tai tietyistä ryhmästä. Tagiavaruus yksinkertaistaa taggaustilannetta, jos käyttäjä luo useita videoita samasta tapahtumasta tai sisällöstä. Se myös helpottaa taggausta mobiililaitteilla, koska käyttäjä voi yksinkertaisesti valita tagit ohjelman tarjoamista vaihtoehdoista. [2; 3]

2. TIETO SÄHKÖISESSÄ MEDIASSA

Tietoteoria on filosofian osa-alue, joka tutkii tietoa ja sitä mitä voimme tietää. Tieteenfilosofia on laajempi filosofian osa-alue, joka tutkii tieteellisen tiedon perusteita ja pohjaa. Tiedon lajeja voidaan eritellä monella tavalla:

- Empiirinen tieto on aistihavaintoon perustuvaa tietoa ja rationaalinen tieto on järkeen perustuvaa tietoa. Näitä käytetään matematiikassa ja logiikassa.
- Propositionaalinen tieto on väitelauseiden muodossa ilmaistavaa tietoa.
- Tieteellinen tieto on tietoa, joka voidaan todistaa tieteellisesti, kun taas arkitieto on ihmisen omiin kokemuksiin perustuvaa tietoa hänen elämänpiiriinsä kuuluvista asioista.
- Intuitiivisella tiedolla on tarkoitettu toisaalta välitöntä tietämistä, kuten esimerkiksi aistitietoa, mutta myös sisäistä intuitiota, välitöntä ymmärrystä. Monet keksinnöt ja tieteelliset oivallukset perustuvat intuitiiviseen tietoon eli välittömään yhtäkkiseen asian ymmärtämiseen.

[6; 7; 8; 9; 10]

Platonin esittämä, klassisen tiedon määritelmän mukaan tieto on ”hyvin perusteltu tosi uskomus”. Jotta esitetty väite olisi tietoa, sen on Platonin mukaan oltava siis tosi ja perusteltua. Niin sanotussa heikossa tiedon määritelmässä voidaan perustelu jättää pois, eli väite on tietoa, jos se on tosi. Platonin oppilas Aristoteles taas ajatteli tietoa modernin tieteen tavoin. Hänen mukaansa uutta tietoa voidaan tuottaa empiirisesti maailmaa havainnoimalla. René Descartes esitti teoksessaan ”Mietiskelyjä ensimmäisestä filosofiasta”, että ehkä on olemassa jonkinlainen ilkeä kaikkivoipa demoni, pettävä jumala (deus deceptor), joka tuottaa subjektille sen havainnot ja nämä antavat täysin virheellisen kuvan todellisuudesta. Descartes ei siis pitänyt havaintoa luotettavana tietolähteenä subjektista riippumattomasta tietolähteestä, vaan ainoa luotettava tiedon lähde on rationaalinen päättely. [6; 11; 12]

Filosofi Ilkka Niiniluodolla tiedon klassinen määritelmä kuuluu näin:

Tieto on:

- (i) hyvin perusteltu
- (ii) tosi
- (iii) uskomus

Tältä esitykseltä vaaditaan totuutta eli vastaavuutta todellisuuden kanssa. Lisäksi vaaditaan perusteluja siitä, miten tiettyyn käsitykseen on jouduttu; joskushan voidaan puhtaasti arvaamalla osua oikeaan. Tässä tullaan siihen jo Sokrateen esittämään

ajatukseen, että todellinen tieto kohdistuu maailman yleisiin lajiipirteisiin, siihen mikä on maailmassa vakioista, invarianttia. Todellinen tieto koskee luonnonlakeja ja on teoreettista: sen avulla voidaan osua oikeaan vastakin. [13]

Niiniluoto vaihtaa klassisessa määritelmässä "uskomuksen" tilalle "semanttisen informaation", joka siis kuvaa enemmän tai vähemmän tarkasti maailmassa vallitsevia asiantiloja. Tämän jälkeen tiedon määritelmä näyttää seuraavalta:

Tieto on

1. väitelauseiden sisältämää semanttista informaatiota
2. joka täyttää perusteluehdon (i) ja
3. totuusehdon (ii).

Tietoa olisi siis tämän mukaan esimerkiksi kielen lauseissa esitetyt maailmaa koskevat väittämät (maa on pyöreä), joka pystytään perustelemaan vaikkapa kenen tahansa toistettavissa olevien tutkimusten nojalla ja joka nojaa parhaisiin käytettävissä oleviin luonnontieteellisiin teorioihin. Tärkeintä kaikesta on kuitenkin Niiniluodon mukaan se, että väittämä on totta eli vastaa todellisuudessa vallitsevaa asiantilaa. [13]

Tämä työ ei ota kantaa onko niin sanottu tieto tosi. Tässä työssä tieto on intuitiivista tietoa ja liittyy siis ihmisen omaan kokemukseen. Tieto on ihmisen ymmärtämään ympäristöön tai johonkin asiaan mediassa liittyvää tietoa tässä työssä. Tärkeintä on saada käyttäjältä realistisia vastauksia ja niiden saamiseen on kaikkien ymmärrettävä tietoa jollain asteella.

2.1. Konteksti- ja metatieto

Ihminen osaa luontaisesti aistia monia kontekstiin liittyviä asioita ja mukauttaa toimintaansa sen mukaisesti. Ilman ymmärrystä kontekstista ihminen toimisi suorastaan järjettömällä tavalla. Karkeimmillaan tämä tarkoittaisi, että ihminen ei tuntisi kipua tai kylmyyttä eikä ymmärtäisi miten näihin tuntemuksiin tulisi reagoida. Sosiaalisen kontekstin hahmottaminen ja sen mukaan sopivasti toiminen, onkin jo selvästi monimutkaisempaa ja vaatii paikallisen kulttuurin tuntemusta. [14]

Kontekstia käytetään määrittelemään esimerkiksi kuvien sisältöä. Kuvien sisältö taas antaa viitteet kontekstin luomiselle. Tutkimuksen mukaan digitaalisten kuvien ottaminen kasvaa noin 24 prosenttia vuodessa [15]. Tämä tutkimus sisältää ainoastaan digitaalikuvat eikä ota kantaa muiden sisältöjen kasvuun, joita voi olla muun muassa teksti, ääni ja video. Jotta ihminen voi järkevästi käsitellä ottamiaan kuvia, tulisi hänen olla tietoinen, edes jollain tasolla, kuvien sisällöstä. Sisällön tunnistamisessa auttaa metatieto, joka voi yksinkertaisimmillaan toimia tageina. [16]

Jotta kontekstia voisi käyttää tehokkaasti, tulisi ymmärtää mitä konteksti on ja miten sitä voi käyttää. Tarkan määrittelyn antaminen kontekstitiedolle on hankalaa. Monet ihmiset ymmärtävät mitä konteksti on, mutta eivät silti osaa selittää sitä muille. Parhaimman määrittelyn kontekstitiedolle saa käyttämällä tutkijoiden määrittelyksiä. [17]

Kontekstitieto on siis tiettyyn tapahtumaan liittyvää informaatiota.

Kontekstitietoa ei voi määritellä absoluuttisesti, koska kontekstitiedolla tarkoitetaan oleellisia asioita tapahtuman ympärillä. Näitä asioita ei pystytä määrittämään tarkkaan. [17] Schilit et al. määrittelevät kontekstin kolmeksi tärkeimmäksi näkökulmaksi sijainnin, seuran ja lähellä olevat resurssit. Kontekstiin kuuluu heidän mukaansa myös monia muita asioita, muun muassa: valoisuus, melutaso, Internet-yhteys ja sosiaalinen tilanne. [18] Sellaiset määrittelyt, jotka määrittelevät kontekstin esimerkein on yleensä vaikea toteuttaa. Kun haluamme määrittää kuuluuko jonkin tyypin informaatio kontekstitietoon, mitä ei ole määrittelyssä mainittu, ei ole helppoa päätellä miten kuuluu tällöin toimia. Kontekstin selventämiseen on siis hyvä käyttää synonyymeja kuten edellä. Siinä mielessä määritelmä on kuitenkin hankala, että sitä on vaikea soveltaa todellisuudessa. [17] Kontekstitieto voidaan jakaa kahteen ryhmään, jotka ovat inhimilliset tekijät ja fyysinen ympäristö. Inhimilliset tekijät voidaan jakaa edelleen kolmeen kategoriaan: tieto käyttäjästä, sosiaalinen ympäristö ja tieto käyttäjän tehtävistä. Myös fyysinen ympäristö voidaan jakaa kolmeen kategoriaan: olosuhteet, infrastruktuuri ja sijainti [19]. Näiden lisäksi aika ja luonnollinen konteksti ovat tärkeitä. Schilit et al. ja Pascoe ovat omassa määritelmässään päätyneet samantyyliiseen ratkaisuun, kuin MoViE. Schilit et al. mukaan, tärkeitä näkökulmia kontekstissa ovat missä olet, kenen kanssa olet ja mitä resursseja on lähellä. Pascoe määrittelee kontekstin fyysisiin ja käsitteellisiin alaryhmiin, joissa ollaan kiinnostuneita tietystä yksiköstä. [17, katso 20] Tämä saattaa kuitenkin olla hieman liian tarkka kuvaus.

On otettava myös huomioon, että tiedon välitykseen liittyy oleellisesti kaksi kontekstin osa-aluetta. Nämä ovat tiedon konteksti ja tiedontarvitsijan konteksti. Kun nämä kohtaavat on tiedon arviointi ja ymmärtäminen helppoa. Tällainen tapahtuu kuitenkin yleensä vain, kun käytetään omia ideoita. Kontekstien erotessa toisistaan paljon, on tärkeä löytää välineitä, miten tieto saadaan välittymään tehokkaasti kontekstista toiseen. Tiedontarvitsijan kontekstin mallintaminen onkin käyttäjän mallintamisen yleisimpiä muotoja. Käyttäjän- ja kontekstin mallintamisella on eri päämäärät. Tämän työn kannalta tärkeämpi on käyttäjän mallintaminen. Tässä kone pyritään saaman ymmärtämään käyttäjää. Toisessa mallintamisessa pyritään saada käyttäjät ymmärtämään toisiaan. Tähän ei kuitenkaan olla työssä otettu kantaa. [21]

Metatieto on tietoa tiedosta, sen avulla kuvaillaan tietovarantojen sisältöä. Metatieto tarjoaa keinon digitaalisten tietojen indeksointia, pääsyä, säilytystä ja etsimistä varten. Digitaalisen tiedon määrä, mikä on saatavilla sähköisessä verkossa, on luonut pakottavan tarpeen standardeille, jotka auttavat sen paikantamisessa, lataamisessa ja hallinnoimisessa. [16, katso 22] Esimerkiksi tietokonetta käytettäessä, kun määritellään tiedostonimi tiedostojärjestelmässä, kuvaa tiedostopolku metatietoa. Perinteisten tiedostojärjestelmien "metatietojärjestelmä" perustuu yleensä kuvaaviin tiedostonimiin ja hakemistorakenteisiin. Metatieto pitää pystyä tallentamaan, käsittelemään ja hyödyntämään sisältöyksiköinä. Sisältöyksiköt ovat organisoidusti tunnistettavaa sisältöä. Metatieto on siis tietovarantoon liittyvää tietoa, joka on tarkoitettu tunnistamaan tietovarannon käyttöä. Metatiedon voi luokitella kolmeen kategoriaan:

- Semanttinen metatieto, joka on sisällön merkitystä kuvaava tieto.
- Rakennemetatieto, joka on sisältöyksikön fyysistä rakennetta, loogista rakennetta tai sisällön kieltä kuvaavaa tietoa.
- Kontekstuaalinen metatieto, joka kuvaa sisältöyksiköiden ympäristöä tietyssä tilanteessa.

Kontekstuaaliseen metatietoon voidaan helposti yhdistää myös kontekstitieto [23]. Kontekstitieto on siis samantyyppistä tietoa kuin metatieto. Edellä esitetty metatiedon luokittelu ei kuitenkaan ole ainoa metatiedon määrittely tapa. Metatieto voidaan tilanteesta riippuen määritellä hyvinkin monella tavalla. Gilliland-Swetland (2000) on kategorisoinut metatietoa esimerkiksi sisällön hallitsemisen näkökulmasta. Tämä määrittely jakaa metatiedon viiteen kategoriaan: hallinnollinen -, kuvaileva -, varjeltu -, tekninen - ja käyttö metatieto. Tämä määrittely on kuitenkin hyvin rajallinen suuremman tutkiskelun kannalta. [16, katso 24; 25; 26]

Anind K. Dey:n artikkelissa Understanding and using context on määritelmä, joka on tämän työn kannalta hyvinkin lähellä haluttua. Sen mukaan konteksti on mitä tahansa informaatiota, jota voi käyttää määrittelemään tilanteen yksiköt. Yksikkö on ihminen, paikka tai esine, mikä on relevanttia vuorovaikutukseen käyttäjän ja sovelluksen välillä, mukaan lukien käyttäjä ja sovellus itse. [17]

Tässä työssä käytetään automaattisesti saatavaa kontekstitietoa apuna tagien tietokantahauissa. Tämän työn tärkein automaattinen kontekstitieto on paikkatieto. Myös monet tutkijat, kuten esimerkiksi Dey et al., Schilit ja Theimer, pitävät paikkatietoa erittäin tärkeänä kontekstitietona [27]. Näiden avulla pyritään saamaan mahdollisimman realistisia tuloksia. Toisin sanoen MoViE -palvelusta on pyritty saamaan mahdollisimman hyvä kontekstitietoinen järjestelmä. Kontekstitietoiset järjestelmät ovat yleensä tehokkaita, koska ne pyrkivät ymmärtämään käyttäjän ympäristöä. Dey ja Abowd ovat määrittelleet kontekstitietoisien järjestelmän seuraavasti: järjestelmä on kontekstitietoinen, jos se käyttää kontekstia tuottamaan olennaista informaatiota ja/tai palveluja käyttäjälle. [28] Ideaalitapauksessa käyttäjän ei tarvitse kontekstitietoisissa palveluissa havainnoida ja välittää palvelulle yhtä paljon tietoa, kuin sellaisissa palveluissa, jotka eivät ole kontekstitietoisia. Tästä päästään ongelmaan, jota tässä työssä ei ole otettu huomioon. Kontekstitieto on aina jollaintavalla henkilökohtaista ja kaikki eivät halua antaa tällaista tietoa käsiteltäväksi ja jaettavaksi. [14; 29]

2.2. Tagi

Käyttäjien mediaan lisäämiä avainsanoja kutsutaan tageiksi. Taggaaminen on yleinen tapa lisätä informaatiota varsinkin sosiaalisen median sovelluksissa [30]. Tagit ovat yksinkertaisia metadata attribuutteja [16]. Avainsanojen lisäämistä puolestaan kutsutaan taggaamiseksi. Koska taggaaminen on vapaamuotoista, voivat avainsanat olla millaisia tahansa. Yleensä tagit ovat asiaa kuvaavia tai muulla tavalla asiasta tietoa lisääviä. Tagit

Vaikkakin taggaus antaa käyttäjälle vapaat kädet valita mitä ja miten taggaa, voidaan sitä silti rajoittaa erilaisin sovelluksin. Eräillä Internet-sivustoilla paheksuvat sanat on poistettu mahdollisten tagien joukosta. Esimerkiksi Amazon.com sivustolla ei voi käyttää tageina kirosanoja. Hyvin käytettynä tagit ovat kuitenkin ihmiselle hyödyllinen väline jäljittämään, jakamaan ja etsimään tietoa. [31]

Tagit ovat yleistyneet huomattavasti lähivuosien aikana, koska ne palvelevat monessa mielessä eri käyttäjiä. Lisäksi ne ovat helppoja ja niiden käyttäminen on vaivatonta. Käytössä ne ovat yksinkertaisia, koska niiden lisäämiseen vaaditaan ainoastaan kenttä, johon tieto syötetään. Tagit ovat myös joustavia, eli niitä voi käyttää kuvailemaan mitä tahansa. Oma tagikirjasto voi kasvaa helposti melko kattavaksi, kuitenkin kokonaisuutta on helppo hallita. [31]

Tageja ja taggaamista ei ole mielekästä ajatella yhtenä konseptina. Taggaaminen on enemmänkin kontekstisidonnaista, monimuotoista sekä moniselitteistä. Taggaamiseen vaikuttaa mm. sosiaalinen konteksti, käyttöliittymä ja järjestelmän avoimuus. Taggaaminen tulisi suunnitella siten, että käyttäjällä on mahdollisuus päättää ja kontrolloida itse, miten ja milloin tageja käytetään. [31; 30]

Tagien liittäminen voidaan myös määritellä erilaisien sääntöjen avulla. Semanttinen web lähestyy tiedon semanttista esittämistä täsmällisten määrittelyjen, standardien ja yhteisesti sovittujen käsitteistöjen kautta [33]. Semanttinen web toimii koneennäkökulmasta tiedonluokitteluun, kun taas folksonomia toimii käyttäjännäkökulmasta. Vander Wal on määritellyt folksonomian olevan informaation luokittelukäytäntö. Ihmiset voivat liittää erilaisiin objekteihin vapaasti asiasanoja, jotta ne ovat löydettävissä myöhemmin uudelleen. Folksonomian suurimpia heikkouksia on termien päällekkäisyydet ja epätarkkakäyttö sekä useasta sanasta koostuvat tagit ja niiden erilaiset kirjoitusasut. Folksonomiassa lisättyjä asiasanoja kutsutaan tageiksi. Folksonomian ja semanttisen webin olennaisin ero on avainsanojen vapaa valinta ontologian (Ontologia on teoria olemassa olevasta, yhteisesti sovitusta muodollisesta tietokonelukuisesta käsitteistöstä [34]) määrittelemien termien sijaan. Folksonomia on suunnattu juuri käyttäjille, koska koneiden on vaikea ymmärtää mitä ihminen eri tilanteissa tarkoittaa. Folksonomian tiukempi versio on taksonomia. Tieteelliset luokittelujärjestelmät perustuvat perinteiseen kategorisointiin eli taksonomiaan. Taksonomiassa ei lisätä pelkkää avainsanaa vaan sille on tehtävä järkevä polku, miten sanaan on päästy. Taksonomia polveutuu semanttisesta webistä, jonka ansiosta se onkin koneille sopivampi käytäntö. [16; 35; 36; 37; 38; 39]

2.2.1. Manuaalinen tagi

Manuaalinen tagi tarkoittaa käyttäjän itse lisäämää tagia. Se voi olla mihin tahansa mediaan lisättävä etsimistä helpottava avainsana. Manuaalista tagia ei välttämättä käytetä yhtä paljoa, jos sovellukseen ei ole erikseen merkitty paikkaa, johon sen voisi lisätä. Toinen huono puoli manuaalisessa tagissa on käyttäjän vaihteleva oma kiinnostus

ja kokemus taggaamisesta. Toisinaan manuaalisen tagin lisääminen sovelluksissa voi olla myös aikaa vievää ja siksi turhauttavaa. Voi olla, että käyttäjä haluaisi lisätä tageja tiedostoihin tai kuviin, mutta ei keksi hyviä sanoja. Tällaisiin tilanteisiin auttaisi, jos käyttäjälle ehdoteltaisiin valmiiksi tageja parametrien avulla. Tämä ei pelkästään auttaisi käyttäjää vaan myös nopeuttaisi toimintaa huomattavasti. Edellyttäen, että syötteenä tulisi ”järkeviä” tageja. [31; 32]

Tagit ovat alkaneet yleistyä myös erilaisissa web-palveluissa. Esimerkiksi Facebook mahdollistaa ihmisten tai esineiden taggaamisen kuvaan, mikä helpottaa tiettyjen kuvien etsintää ja antaa niille lisäarvoa. Tageja käytetään pääasiassa tietojen etsimiseen. Tätä käyttävät hyväksi erilaisten hakukoneiden avulla toimivat mm. erilaiset myyntisivustot, kuvien selaukseen tarkoitetut sivut sekä blogit. Tageja käytetään kuitenkin niiden potentiaaliin nähden vielä melko vähän.

Kukaan ei taggaa pelkästä taggaamisen riemusta. Taggaaminen ei siis ole itsessään päämäärä vaan työkalu. Taggaaminen on kuitenkin vielä monelle varsin vieras konsepti. Toiset ymmärtävät tagit, mutta eivät ole vielä saaneet niistä riittävää hyötyä jokapäiväiseen käyttöön. Tagien käytön helpottaminen voisi monilla laskea kynnystä niiden käyttöön. Tagit kuitenkin helpottavat myöhemmin tiedon etsimistä huomattavasti. Tulevaisuudessa tagien käyttö yleistyne, kun tietomäärät kasvavat ja hakujen tehokkuus korostuu. Tagien käyttö helpottaa tietojen etsimistä niin paljon, että tagit tulee vielä jokapäiväiseen käyttöön. Samasta syystä myös käyttöjärjestelmiin kehitettäneen erilaisia tageihin perustuvia hakukoneita. [32; 40]

2.2.2. Automaattinen tagi

Ongelma tagien käytössä on usein se, että kaikki ihmiset eivät viitsi tagata tietojaan, vaikka siitä olisikin hyötyä heille. Tämän voi huomata kaikista taggauksen mahdollistavista sovelluksista, joita käyttää hyväkseen vain osa ihmisistä. Tämän ongelman pysyisi kiertämään siten, että sovellus ehdottaisi joitain ”järkeviä” tagisanoja automaattisesti, mikä taas vähentäisi käyttäjän taakkaa. [31]

Automaattinen tagi tarkoittaa tagia, joka tulee automaattisesti syötteenä ohjelman avulla. Jotta tällainen sovellus saataisiin toteutettua järkevästi, vaatisi se käyttäjästä tietoa, esimerkiksi paikasta tai ajasta. Tietojen perusteella voitaisiin ehdottaa samassa paikassa tehtyjä tageja, tai vastaavasti samaan aikaan muualla tehtyjä tageja. Sovellus toimisi periaatteessa siten, että käyttäjä ottaa esimerkiksi kuvan ja lataa tämän palvelimelle. Palvelin tarkistaa, keneltä kuva tulee, ja ottaa sen mukaan tietokannastaan käyttäjän sinne ennestään syöttämiä tageja. Ohjelma toimisi käyttäjältä huomaamattomasti ja vaivattomasti, koska kyselyt tapahtuisivat tausta-ajona. Kun kyselyt olisivat valmiit, syötteenä tulisi lista tageista, joista käyttäjä voisi valita sopivimmat ja säästyisi samalla turhalta kirjoittamiselta. Ongelma tällaisessa on, ”Mikä on sopiva ehdotusten määrä?”, koska suurin osa käyttäjistä ei kuitenkaan viitsi käydä loputonta listaa läpi. [31; 32]

Automaattisesti tuleva tagi voi olla myös sellainen, joka tunnistaa kuvasta esineen tai asian ääriiviivat ja etsii niiden mukaan tietokannan ja palvelimen avulla monimutkaisilla laskuilla tagisanan. Tällaisia sovelluksia ei haastavuuden vuoksi ole kehitetty moniakaan. Paljon helpompi on luoda esimerkiksi sovellus, joka tunnistaa montako ihmistä kuvassa on eikä niinkään onko kuvassa joku tietty henkilö [16]. Lisäksi kyseinen sovellus ei kuitenkaan palvele tageja riittävän paljoa ollakseen tarpeellinen, koska tagit voivat olla myös muuta, kuin kuvassa näkyvä konkreettinen esine. Lisäksi videoissa vastaavanlainen menettely ei ole kovinkaan kannattavaa, sillä yhden videon aikana tallentuu monta kuvaa (Nokia 5800 XpressMusic tallentaa 30 kuvaa sekunnissa parhaalla laadulla [41]). Siitä johtuen kuvien hahmojen laskennallinen tunnistus olisi järjestelmälle raskasta.

2.2.3. Tagikirjasto

Tagikirjastolla tarkoitetaan automaattisten tagien joukkoa, joka kehittyy automaattisesti sovelluksen eliniän aikana. Tagikirjasto voi olla jokaisen henkilökohtainen tai yleinen. Se voi perustua käyttäjän omiin syötteisiin tai palvelimella tapahtuviin kyselyihin ja vasteisiin. Tagikirjasto sisältää tageja, joita palvelimelle on syötetty aikaisemmin tai mitä palvelin hakee muualta esim. Internetin hakukoneista. Tagikirjasto voi olla samantyylinen kuin tagipilvi. Kirjasto sisältää sanoja, joista toiset ovat toisia merkitsevämpiä. Jokaisen henkilökohtainen tagikirjasto kasvaa sitä mukaa, kun syötteitä, eli tagisanoja kertyy lisää. Tagikirjasto voi olla myös yleinen, eli syötteenä voi tulla myös muiden ehdottamia tageja. Tähän vaaditaan järjestelmältä jo jonkinlaista tekoälyä. [31]

Yhteinen tagikirjasto toimii yleensä jonkinlaisen tekoälyn avulla. Se voidaan rakentaa toimimaan palvelimeen tai vaikkapa matkapuhelimeen. Palvelimen avulla toimiva tagikirjasto voi olla yleinen esimerkiksi siten, että se ehdottaa tageja paikan, ajan tai sään mukaan. Nämä kyselyt ja vasteet tulevat yleisimmin tietokantaan tehtyjen kyselyiden perusteella. Kyselyitä voi olla myös monia. Edelläkuvatus menettelyn avulla saadaan kirjastosta hyödyllinen ja käyttäjäystävällinen. Tagikirjasto ei siis yleensä ole valmiiksi rakennettu, vaan se kehittyy sitä mukaa, kun käyttäjät lisäävät omia tagejaan.

Tagikirjaston käytettävyyttä voidaan pohtia monelta kannalta. Kannattaako tehdä sellaista kirjastoa, joka antaa ”turhia” tagisanoja. Kaikki eivät kuitenkaan käytä samoja sanoja taggaukseen. Ihmiset näkevät eri asioita tagatessaan videon tai kuvan. Kuitenkin suurin osa samassa ympäristössä olevista ihmisistä käyttää samantyyppisiä tagisanoja. Ihmiset yrittävät yleisesti nähdä jotain suurempaa kokonaisuutta, tai pieniä yksityiskohtia, joita voi käyttää tagi sanoina. Tämä käy ilmi myös tutkimuksesta, joka toteutettiin osana diplomityötä. [Liite 1] Tutkimuksesta kerrotaan tarkemmin luvussa 5.1.1. Tagikirjaston hyödyllisyys. [31]

Tagikirjastoa käytettäessä on otettava myös huomioon, halutaanko käyttää henkilökohtaisia tageja, joita ei jaettaisi muille. Tässä työssä on päätetty käyttää

tagikirjastoa yleisenä ja kaikki tagit, joita kirjastoon syötetään, tulevat kirjaston käyttöön. Tämä johtuu osin ohjelman pienestä käytöstä ja tarpeesta saada kattavampi tagikirjasto. Toisaalta tageiksi ei yleensä syötetä niin yksityiskohtaisia sanoja, joita muut eivät voisi nähdä. Lisäksi tagit ovat Internet-sivuilla kuitenkin kaikkien nähtävissä.

2.3. Geotagi

Geotaggaaminen on prosessi, jossa lisätään maantieteellinen metatieto erilaisiin medioihin, kuten kuviin, videoihin, Internet-sivustolle tai RSS -syötteisiin (Really Simple Syndication). Kyseinen tieto sisältää yleensä leveysasteen (latitude) ja pituusasteen (longitude) (katso kuva 2.2.). Nämä tiedot ilmoitetaan koordinaatteina. Geotiedot voivat sisältää myös korkeuden, suunnan, tarkkuuden, nopeuden ja paikan nimen. [42, katso 43]

Geotaggaaminen voi auttaa käyttäjiä heidän etsiessä paikkaan perustuvaa tietoa. Käyttäjä voi esimerkiksi etsiä kuvia, jotka on otettu lähellä annettua paikkaa. Tämä tapahtuu syöttämällä geotagihakukoneelle leveys- ja pituusaste ja valitsemalla hakukoneesta kuvien haku. Tietoa, joka sisältää geotagit, voidaan käyttää myös palveluissa, joissa etsitään tietyltä alueelta uutisia, Internet-sivustoja tai muita tietoja, kuten ravintoloita.

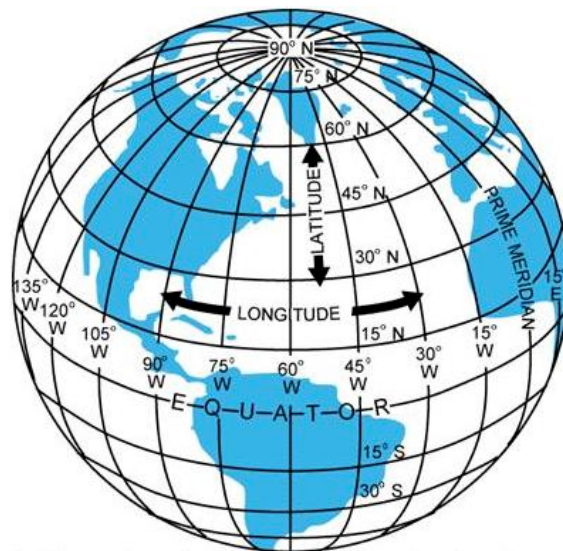
YouTube on suosituimpia sovelluksia, joka mahdollistaa geotagien käytön. Videon lataamisen yhteydessä on mahdollista antaa videolle maantieteellinen sijainti. Sijainnin antamisen jälkeen videot on mahdollista saada näkyville Google Maps palvelussa.

2.3.1. GPS -paikannus

Geotaggaamisessa käytetyistä tekniikoista suosituin on antaa paikkatiedot. Tiedot tulevat GPS -tekniikkaa (Global Positioning System) käyttäen kätevästi, eikä vaadi käyttäjältä mitään jatkotoimenpiteitä. GPS on Yhdysvaltain puolustushallinnon ylläpitämä satelliittipaikannusjärjestelmä, jota kutsutaan myös nimellä Navstar GPS. Nykyään voidaan käyttää myös termiä GNSS. GNSS voi tarkoittaa useita eri satelliittipaikannusjärjestelmiä, joista merkittävimmät ovat Navstar GPS, GLONASS ja Galileo. Näistä kaksi ensimmäistä ovat jo käytettävissä. Navstar GPS:n kehitys on aloitettu 1970-luvulla ja siihen kuuluu nykyään 24 satelliittia, jotka kiertävät maapalloa. Täydellinen 24 satelliitin järjestelmä saatiin valmiiksi huhtikuussa 1994. Järjestelmä on tarkoitettu sekä siviileille että armeijan käyttöön. Satelliittien kiertoradat on järjestetty siten, että jokaisessa pisteessä maapallon pinnalla on jatkuvasti näkyvissä vähintään neljä satelliittia. Käytännössä satelliitteja näkyy kuitenkin aina enemmän kuin neljä, jolloin paikannusvastaanotin voi valita sopivimmat satelliitit käyttöönsä. Paikannukseen voidaan myös käyttää useampia satelliitteja, mutta neljällä satelliitilla saadaan jo

toteutettua kolmiulotteinen paikannus (leveys, pituus ja korkeus). Paikannuksen perustana on kolmiomittaus satelliitista, joka toteutetaan radiosanomien kulkuajan mittauksella satelliittien ja vastaanottimen välillä, kun satelliitin sijainti tunnetaan. Mittauksia varten tarvitaan virheiden minimoimiseksi erittäin tarkkoja atomikelloja, jotka sijaitsevat satelliiteissa. [44; 45; 46; 47]

GPS:n tarkkuus on muutamia metrejä. GPS -paikannuksessa saattaa syntyä virheitä useissa vaiheissa satelliitista vastaanottimeen. Satelliittien radat muuttuvat mm. maapallon vetovoiman ja pyörimisliikkeen vaikutuksesta ja näin syntyy ratavirheitä. Kellovirheitä esiintyy niin satelliitissa kuin vastaanottimessakin. Troposfääri ja ionosfääri vaikuttavat signaalin kulkunopeuteen, ja tämä aiheuttaa luonnollisesti paikannusvirhettä. Näiden lisäksi paikannusvirhettä voi aiheuttaa tahallinen häirintä tai käyttäjän tekemät virheet. [47; 48]



Kuva 2.2. Maapallon pituus- ja leveysasteet [49]

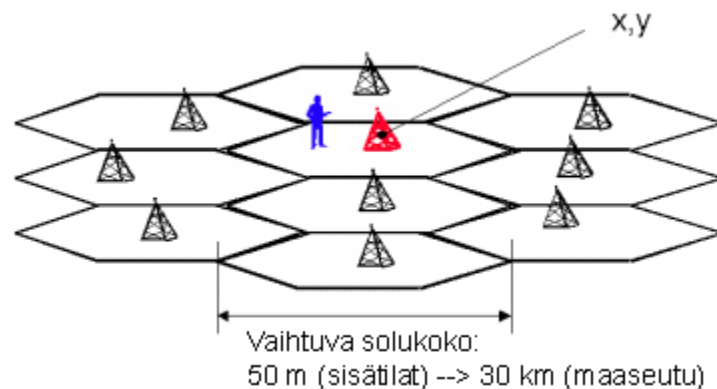
GPS:n avulla saatavia paikkatietoja voidaan käyttää esimerkiksi geokoodaamiseen. Geokoodaaminen ei ole aivan samanlaista kuin geotaggaaminen. Geokoodaaminen perustuu paikan antamiseen sekä pituus- että leveysasteena. Vasteena tulevat paikkatiedot, eli maa, kaupunki ja osoite. Tällaisia sovelluksia löytyy Internetistä jo melko runsaasti. Eräs tämän tyylinen palvelu on Geonames.org sivustolla. Käyttäjä syöttää sivuston osoiteriville halutun paikan leveys- ja pituusasteen ja saa syötteen xml-muodossa. Syötteestä saa selville syötettyjen koordinaattien perusteella paikan. [50: 51; 52]

2.3.2. Solupaikannus

GPS -tiedon lisäksi voidaan paikan määrittämiseen käyttää myös GSM -solutietoja. Solutiedon avulla määritelty paikka ei kuitenkaan ole yhtä tarkka, kuin GPS:n avulla

määritelty paikka. GSM -solun avulla saadaan tiedot kaupungin tarkkuudella. Solupaikannus on yksinkertaisimpia menetelmiä paikantaa matkapuhelin GSM -verkossa. Kuva 2.3. esittää GSM -soluverkkoa. [1]

GSM -verkkoa kutsutaan soluverkoksi. Tämä johtuu siitä, että jokainen verkon tukiasema peittää tietyn maantieteellisen alueen. Aina yhden tukiaseman peittoaluetta kutsutaan soluksi. Solupaikannus perustuu matkapuhelimenverkon solun tunnistukseen. Matkapuhelimen sijainti voidaan määrittää solun koon tarkkuudella, jos tukiaseman sijainnin koordinaatit tunnetaan. Solujen koot voivat vaihdella 50 metrillä (sisätilat) 30 kilometriin (maaseutu). [1]



Kuva 2.3. GSM -solut muodostuvat tukiasemien avulla.

Solupaikannuksen ongelmana on riippuvuus solun koosta. Esim. maaseuduilla on erittäin epätarkka paikannus, johtuen matkapuhelinverkon tukiasemien pidemmistä etäisyyksistä. Kaupunkialueella sen sijaan päästään tilanteesta riippuen melko hyvään paikannustarkkuuteen (200-500m). Huomioitavaa kuitenkin on, että palveleva tukiasema ei välttämättä ole lähin tukiasema. Tämä johtuu tukiasemien suurehkoista määrästä. Etuna solupaikannuksessa on kuitenkin sen helppous. Solupaikannuksen tarkkuutta voidaan parantaa halutessa vielä ajastusennakon (Timing Advance, TA) avulla. [1; 53]

2.3.3. Verkkoavusteinen paikannus

Nykyään on mahdollista avustaa satelliittivastaanottimen paikanmäärittystä myös verkkoyhteyden avulla. Toisinaan satelliittien lähettämä signaali voi olla heikko. Signaalin havaitseminen esimerkiksi sisätiloissa voi olla hankalaa, kun suoraa näköyhteyttä satelliitteihin ei voida saada aikaan. Signaalin saatavuus voidaan kuitenkin ennustaa, jos vastaanottimen paikka tiedetään karkeasti. Tällaisessa tapauksessa paikannuslaitteelle pyritään toimittamaan erilaisia tietoja, jonka avulla signaali voidaan pyrkiä suodattamaan esiin kohinasta erityisillä menetelmillä. Kyseinen toiminta edellyttää kuitenkin laskentatehoa, jota varten vaaditaan tietokonetta. Tämä toteutetaan siirtämällä tarvittavat tiedot tietoliikenneverkon kautta erityisen paikannuspalvelimen tehtäväksi. [47]

Tämän toimenpiteen avulla pystytään sijainti määrittämään entistä nopeammin sekä myös melko tarkasti. Sijainti pystytään myös määrittelemään ahtaissa kaupunkitiloissa sekä kevytrakenteisissa rakennuksissa. [54] Tällaista sovellusta käytetään varsinkin uusimmissa matkapuhelin malleissa. Sen tarkoituksena on parantaa matkapuhelimen akun kestoa sekä parantaa ja nopeuttaa paikannustarkkuutta. Sovellusta kutsutaan nimellä A-GPS (Assisted GPS). A-GPS on myös eräs standardoiduista paikannusmenetelmistä 3G –verkossa. [28; 47; 55; 56]

Satelliittipaikannusta voidaan avustaa myös anturien tuottamalla informaatiolla. Anturien avulla voidaan sijaintitietoa säilyttäen laskea sijainti merkintälaskun avulla, vaikka satelliittien signaalien paikannus ei onnistuisikaan. Tällainen on hyödyllinen esimerkiksi ajettaessa tunnelin läpi. [54]

3. TIEDON LIITTÄMINEN SÄHKÖISEEN- JA MOBIILI MEDIAAN

Erilaista tietoa voidaan liittää sähköiseen mediaan, esimerkiksi kuvaan tai videoon, nykyään melko helposti. Sovelluksia löytyy Internetistä paljon. Monissa sovelluksissa tiedon liittäminen tarkoittaa tagien ja selitetekstin liittämistä kuvaan. Tällaiset sovellukset ovat hyviä siinä mielessä, että myös ulkopuoliset kuvien tai videoiden katselijat saavat niistä enemmän informaatiota. Kaikki eivät kuitenkaan ole kiinnostuneita näiden lisätietojen lisäämisestä, koska näkevät sen hankalaksi ja aikaavieväksi työksi.

Lisätietojen ansiosta on kuitenkin mahdollisuus linkittää kuvia tai videoita moniin eri asioihin. Erityisesti lisätiedot toimivat, kun suoritetaan jonkinlaisia etsintöjä tai linkityksiä. Tietomäärien kasvaessa tämä on erittäin hyödyllinen ominaisuus.

3.1. Tiedon liittäminen kuvaan

Eräs suosituimpia yhteisöjä, joissa voi tagat kuvia ja liittää niihin informatiota on Facebook. Sovelluksen tietää lähes kaikki ympäri maailmaa ja sen nimi herättää heti mielipiteitä ihmisissä. Suomessa sivusto on ollut tunnettu vuodesta 2007 lähtien. Sen käyttäjäkunta on monipuolista, niin iältä kuin käyttötavoilta. Sen suuren tagien liittämismahdollisuuksien vuoksi on se hyvä esimerkki tässä käytötapauksessa.

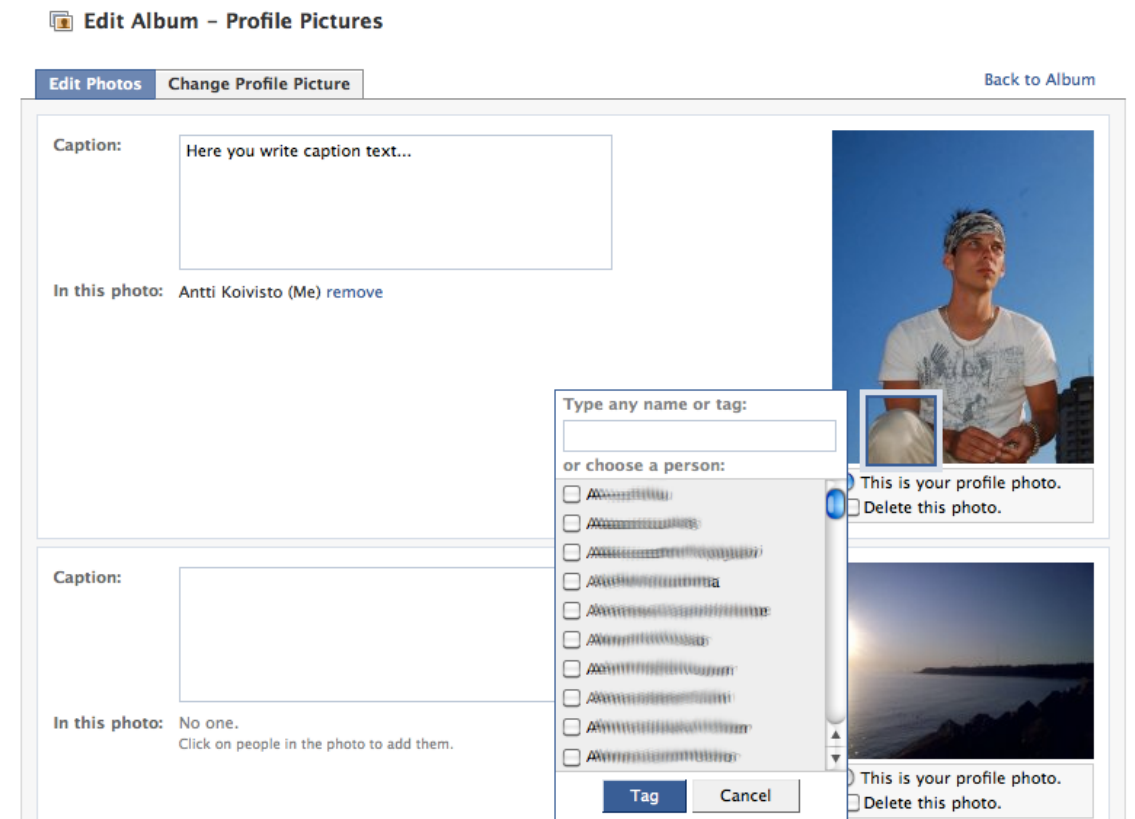
Facebook on Internetissä toimiva sosiaalisen median yhteisöpalvelu. Sivusto tarjoaa käyttäjille paljon eri mahdollisuuksia, mm. käyttäjäprofiilin ja blogin luomiseen, sekä yhteydenpitoon ystävien kanssa ympäri maailman. Facebookia ylläpitää samaista nimeä kantava yritys Yhdysvalloissa, Palo Altossa Kalifornian osavaltiossa. Facebook on perustettu 4. helmikuuta 2004 Harvardin yliopiston kampuksella Mark Zuckerberg toimesta. Facebookilla on yli 900 työntekijää. Facebookiin on liittynyt maailmanlaajuisesti jo yli 350 miljoonaa ihmistä. [57]

Lokakuussa 2007 Microsoft osti 1,6 %:n osuuden Facebookista. Kaupan arvo oli 240 miljoonaa dollaria, eli noin 169 miljoonaa euroa. Tämän mukaan Facebookin arvo on noin 10,5 miljardia euroa. [58]

Facebook oli suunniteltu alun perin oman yliopistonsa sisäiseen käyttöön, mutta kysynnästä johtuen se laajeni vähitellen. Lopulta vuonna 2006 siitä tuli täysin avoin. Suomessa Facebook alkoi levitä varsinaisesti vuoden 2007 lokakuussa. Tuolloin aktiivisia käyttäjiä oli kokonaisuudessaan jo 58 miljoonaa. [58; 59]

Yksi Facebookin suosituimmista ominaisuuksista on kuvien lataaminen ja

tilaviestien päivittäminen. Kuville voidaan tehdä hyödyllisiä asioita linkitettävyyden ja informaation lisäämisen kannalta.



Kuva 3.1. Kontekstitiedon lisääminen kuvaan Facebookissa.

Palvelussa voidaan kuvaan linkittää ihmiset, minkä ansiosta kuvat linkittyvät niin ystäville, kuin itse linkitetylle ihmiselle (katso kuva 3.1.). Tällainen sovellus on melko hauska, joskaan ei välttämättä hyödyllinen. Tästä johtuen on Facebookissa käyttäjiä, jotka vain tarkastelevat muita, eivätkä käytä sitä niin kuin ”olisi tarkoitettu”. Myös sellaisia käyttäjiä on paljon, jotka eivät halua itseään tagatuksi. Tähän Facebook ei ole ottanut mitenkään kantaa. Ainoa tapa on selata kaikki kuvat erikseen ja poistaa yksitellen jokainen tagi kaikista kuvista.

Kuviin voidaan myös lisätä kommentteja, jonka ansiosta niihin tulee mahdollisesti lisää informaatiota. Kuviin lisättävät kommentit ovat vapaamuotoisia, mikä tarkoittaa, ettei kuviin lisättävät kommentit välttämättä lisää kuvan arvoa mitenkään. Kommenttien lisääminen ja ihmisten taggaaminen kuviin on kuitenkin erittäin suosittua Facebookin sivustoilla. Se tekee kuvista mielekkäämpiä katsella ja antaa monille käyttäjille tietoa, jota he eivät välttämättä muuten saisi. Esimerkiksi kuvassa voi olla maininta siitä missä se on otettu, kuka kuvassa on ja milloin kuva on otettu. Tällaisia sovelluksia olisi hyvä kehittää myös muissa ympäristöissä, kuin Facebookissa.

Facebookille on suuren suosion myötä tulossa paljon kilpailijoita. Esimerkiksi Google on ilmoittanut ryhtyvänsä kilpailuun Facebookin kanssa. Myös Suomessa on

vastaavanlainen sovellus, Lyyra, joka on pääasiassa tarkoitettu opiskelijoille. Sen idea on hieman samantyylinen, kuin mistä Facebook lähti liikkeelle. Facebookille on tulossa kilpailua myös erilaisten mikroblogipalveluiden osalta, koska monet käyttäjät eivät ole halukkaita käyttämään sovelluksia, joita Facebook tarjoaa. Toiset ihmiset haluavat vain kertoa, mitä he tekevät tai mitä mielteitä heillä on. Tällaiseen tarkoitukseen on kehitetty mm. Jaiku (suomalainen), Qaiku (suomalainen, korvaa kuihtuvan Jaikon) ja Twitter. Nämä eivät tue kuvien lähettämistä tai mitään muutakaan ylimääräistä, vaan keskittyvät vain lyhyiden tilaviestien päivittämiseen ja niiden kommentoimiseen. Tilaviestit ovat lyhyitä ajatuksia tai ilmoituksia, yleensä käyttäjästä itsestään, mitkä kertovat käyttäjän tai käyttäjän ympäristön sen hetkisestä mielentilasta jotain informaatiota muille. [60; 61; 62]

3.2. Kontekstitedon liittäminen videoon

2000-luvun ensimmäisen vuosikymmenen jälkipuoliskolla on ilmennyt verkkovideoiden huomattavaa suosion kasvua. Tähän on vaikuttanut muun muassa sosiaalikultuuriset seikat ja teknologian kehittyminen. Sosiaalikultuurisilla seikoilla tarkoitetaan videoiden tuloa osaksi elämäntapaa ja arkea. Teknologiassa taas suurin merkitys on ollut kamerapuhelimilla sekä videoiden saatavuudessa tapahtunut kehitys. [63]

YouTube on maailman suurimpia videonlataamispalveluita. Videoon voidaan liittää tageja, niin hakusanoiksi, kuin selitteiksi keskelle videota. Videoita linkitetään ympäri maailmaa. Esimerkiksi uutisten yhteydessä voi olla linkki videoon, joka on ladattavissa YouTubesta. YouTuben suurta suosiota kuvaa muun muassa Samsungin halu liittää YouTube suoraan televisioon. Uusissa LED-TV:issä on mahdollista katsoa YouTuben tarjoamia videoita suoraan televisiosta Internet-yhteyden yli. [64] Tästä syystä Youtube on erinomainen esimerkki videoon liitettävästä informaatiosta.

YouTube on yksityinen osakeyhtiö, joka on perustettu 15. helmikuuta 2005. Perustajina on kolme PayPalin työntekijää: Chad Hurlley, Steven Chen ja Jawed Karim. Youtube on Googlen omistama suosittu videopalvelu, joka on käännetty jo 19 eri kielelle. YouTubessa käyttäjä voi lisätä omia videoita tai katsoa ja ladata muiden käyttäjien lisäämiä videoita. YouTuben toimisto sijaitsee San Brunossa, Kaliforniassa Yhdysvalloissa. YouTube on myös Internetin suosituin suoratoistovideopalvelu. [65; 66] YouTuben videoiden formaattina käytetään Adobe Flash Video-muotoa (.FLV).

Google osti vuoden 2006 loppupuolella YouTuben noin 1,3 miljardilla eurolla. Palvelu ei ollut tehnyt tätä ennen lainkaan voittoa ja Google joutuikin maksamaan 15 miljoonaa dollaria yhtiön velkoja. [67; 68] Siihen on kuitenkin pyritty lisäämään palveluita ja mainoksia. YouTube on myös aloittanut HD -tasaisen videoiden tallentamis- ja katsomismahdollisuuden. [69]

YouTube tunnistaa muutamat videotallennusformaatit, jotka se konvertoi itse .FLV muotoon. Myös MoViE palvelussa on käytetty tällaista videoiden konvertointia,

jotta ne voitaisiin näyttää sivustolla halutulla tavalla. Videot tulevat yleisimmin MoViE – palveluun 3GP-muotoisena. Näitä videoita ei voida näyttää suoraan MoViE:n soittimessa, vaan ne konvertoidaan, eli muunnetaan, sivuston tuettuun muotoon.

The screenshot shows the YouTube video upload interface. At the top, there's a search bar and navigation links like 'Home', 'Videos', 'Subscriptions', 'History', and 'Upload'. Below this, the video title 'mac.mp4' is displayed. The main content area is divided into two columns. The left column contains form fields for 'Video Information' (Title, Description, Tags, Category) and 'Video Thumbnail' (three image choices). The right column shows a video player for 'mac.mp4' with a play button and a progress bar. Below the player is the 'Date and Map' section, which includes a date picker (December 14, 2009), a map location input (geolat=61.4928056430333, geolon=21.81), and a map showing the location in Finland with a red pin and a speech bubble indicating 'This location has been recorded. Altitude: 30 ft'.

Save Changes or cancel

Video Information

Title: mac.mp4

Description: pakkuröö

Tags: testi mac video por

Category: Science & Technology

Video Thumbnail

Choose a video still from above to represent your video in search results and other displays. You can choose a different still image by clicking on it. Note: it can take up to 6 hours for your image to be updated.

Broadcasting and Sharing Options

Privacy

☒ Share your video with the world (Recommended)

☐ Private (Viewable by up to 25 people)

Comments

☒ Allow comments automatically

☐ Allow friends' comments automatically, all others with

Date and Map

Date: December 14, 2009 Today Clear

Map Location: geolat=61.4928056430333 geolon=21.81 Search Clear

Altitude (in ft): 30 Set Clear

This location has been recorded. Altitude: 30 ft

Map Satellite Hybrid

Map data ©2009 Tele Atlas

Kuva 3.2. YouTubeen mahdollisuus liittää erilaista kontekstittietoa videoihin.

Kun käyttäjä lisää videoita YouTubeen, niihin voi liittää erilaisia tageja (katso kuva 3.2.). Videoita voi tämän jälkeen hakea niiden perusteella. Tagien käyttö on siis tässä sovelluksessa erittäin suuressa roolissa. Latausvaiheessa voidaan myös lisätä lyhyt kuvaus videosta sekä otsikko. Tagit ja selitetekstit videosta ovat erittäin tärkeässä roolissa myös hakutoimintojen kannalta. Nykyään verkon käyttäjät ovat erittäin tarkkoja ajankäytöstään. Jos haluttuja tietoja ei löydy, sivusto saatetaan hylätä kokonaan. [63]

YouTube on toiminnaltaan samantyylinen, kuin projektissa toteutettu MoViE sivusto. MoViE – sivusto pyytää samoin käyttäjältä otsikon, kuvauksen ja tagit jokaiselle videolle.

YouTubeissa kaikki eivät pääse lisäämään informaatiota videoon, kuten Facebookissa. Esimerkiksi vain videon lataaja voi laittaa videoon lisäkommentteja, jotka näkyvät, kun videota katsellaan. Tämä on kuitenkin erinomainen toiminto ja lisää videoiden informatiivisuutta paljon. Lisätiedot voidaan asettaa erilaisina informaatiolaatikoina. Vaihtoehtoina on puhekupla, pitkäaikainen selite, joka alkaa esimerkiksi videon alusta ja kestää loppuun asti, sekä laatikko, johon voi lisätä selitteen kuvaamaan videon sisältöä. Kaikkia edellä mainittuja voi vielä korostaa värein ja niistä voi tehdä esimerkiksi linkkejä. Mahdollisuuksia on rajattomasti. Samanlaista sovellusta on kehittänyt myös Viddler. Viddler on luova web-sovellus, jonka avulla voidaan ladata, parannella ja jakaa digitaalisia videoita nopeasti ja helposti selaimen avulla. Viddler on vienyt tagien ja kommenttien lisäystä pidemmälle. Se mahdollistaa kirjautuneilta käyttäjiltä myös muiden videoihin esimerkiksi tagin lisäämisen. Viddler on myös tehnyt tagien ja kommenttien lisäämisen erittäin helpoksi. Ominaisuus on videosoittimessa selkeästi esillä ja sitä on erittäin helppo käyttää, jopa helpompaa kuin Facebookissa. [70]

YouTube mahdollistaa myös ”kotisivujen” tekemisen. Tämä toimii linkittämällä videoita toisiinsa. Eli yksikertaisimmillaan katsotaan videota ja kysytään käyttäjältä esim. oliko video kiinnostava, käyttäjä klikkaa hiirellä videota siinä kohdassa, jossa lukee kyllä ja video alkaa alusta. Tämä toteutetaan luomalla erilaisia linkkejä videoihin. Käyttäjällä on mahdollista luoda videoon laatikoita mistä klikkaamalla pääsee seuraavaan videoon ja niin edespäin. Kun videoita linkittää riittävän paljon ja monipuolisesti saadaan luotua kiinnostavat ”kotisivut”. Tällaisten sivujen tekeminen on kuitenkin melko aikaa vievää. (www.booneoakley.com)

4. TAGI JA KONTEKSTITIETO MOBIILILAITTEESSA

Mobiililaitteilla, kuten matkapuhelimella, saadaan nykyään paljon kontekstitietoa ympäristöstä ja paikasta. Erityisen paljon tietoa saadaan, jos tiedon keräämistä varten on rakennettu palvelinsovellus. Pelkästä automaattisesti mobiililaitteesta lähtevästä tiedosta saisi jo melko kattavan tagiavaruuden, mutta automaattisesti lähetettävää tietoa hyväksikäyttämällä ja vähän älyä lisäämällä saadaan huomattavasti enemmän tietoa tagattavasta mediasta. Toisaalta tällainen vaatii myös käyttäjältä tagin merkityksen ymmärtämistä.

4.1. Sovelluksia

Mobiililaitteilla toimivia kuvan taggauksen mahdollistavia Internet-palveluja on ollut saatavilla jo jonkin aikaa. Näistä tunnetuimpia ovat ZoneTag, Flickr, Jaiku, TagMap ja Facebook. Flickr on verkkopalvelu, jossa käyttäjillä on mahdollisuus julkaista ottamiaan valokuvia ja jakaa ne muiden kanssa sekä arvostella muiden kuvia. Flickr mahdollistaa myös kuvien luokittelun ja niiden hakemisen. Se on myös yksi Internetin suosituimmista kuvapalveluista. Palvelu on maksuton ja siihen on mahdollisuus ladata ilmaiseksi jopa 100 megatavua kuvia kuukaudessa. Kuitenkaan videontaggaamisohjelmistoja ei ole kehitetty mobiililaitteille yhtä ahkerasti. Edellä mainitut ohjelmat perustuvat kuvaan liitettävään kontekstitietoon, yleisimmin tageihin. Ohjelmat ovat kuitenkin saaneet paljon käyttäjiä ympäri maailmaa. Ohjelmistoista ainoastaan ZoneTag tarjoaa käyttäjälle mahdollisuuden käyttää asiakas – ohjelmistoa kuvien ja tagien lataamiseen Internetiin. ZoneTag:ia käytetään liitettäessä kuvia mobiililaitteella Flickr web – sivustolle. ZoneTag toimii siis niin sanottuna rajapintana kuvien lataamisessa Flickr-palveluun. Myös TagMap on kehitetty Flickr:n kanssa yhteensopivaksi. Muut ohjelmistot ovat luoneet web – sivuston, jonka tarkoitus on palvella käyttäjää myös muissa tarkoituksissa kuin kuvien lataamisessa ja taggaamisessa. [71; 72]

Ohjelmien idea on yksinkertainen. Idea perustuu käyttäjän ottamien kuvien lisäämiseen, niiden tagaamiseen ja verkkoon lataamiseen. Suosituin sovelluksista Suomessa on Facebook, joka toimii samalla keskustelusovelluksena. Facebook myös mahdollistaa käyttäjien taggaamien kuvien liittäminen toiselle käyttäjälle heidän omaan

profiiliin. On kuitenkin epäilty Facebookin väistyvän myös Suomessa ja seuraavan tulokkaan olevan Twitter, joka ei ole täysin vastaava toiminnaltaan, kuin Facebook. Keveytensä ansoista se toimii paremmin kännyköissä ja suuren massan käytössä. Tämä viittaa siihen, että ihmiset käyttävät yhä enemmän palveluja mobiililaitteilla. Facebook on kuitenkin vastannut kilpailuun uudella Lite versiollaan, joka on huomattavasti nykyistä versiota kevyempi ja nopeampi. Keveys johtuu lähinnä applikaatioiden pois jättämiseen. [73] Twitter ei myöskään ole saanut vielä suurta suosiota Suomessa. Tuomas Lonka tutki Twitterin suomalaisten käyttäjien määriä. Tutkimuksen tuloksena oli, että suomalaisia käyttäjiä on yli 4000 mutta kuitenkin huomattavasti alle 10 000. [74]

Palvelut, kuten Twitter ja Facebook, ovat yleistyneet paljon, koska niiden käyttäminen on mahdollista tehdä kännykällä missä vaan. Sosiaalisen webin palveluiden suosio on suuressa kasvussa ja tutkimukset kertovat suuren yleisön osallistuvan yhä enemmän sosiaalisen nettisisällön tuottamiseen [35]. Palveluita on helppo käyttää ja niiden antamat mahdollisuudet kiinnostavat myös kehittäjiä. Näiden palveluiden käyttöön ei tarvita muuta kuin Internet-yhteys ja selain. Miltei kaikissa kännyköissä on kamera ja mahdollisuus jonkinlaiseen Internet-yhteyteen. Jos kännykässä ei ole valmiutta tähän, on kaikki mahdollista tehdä myös tietokoneella.

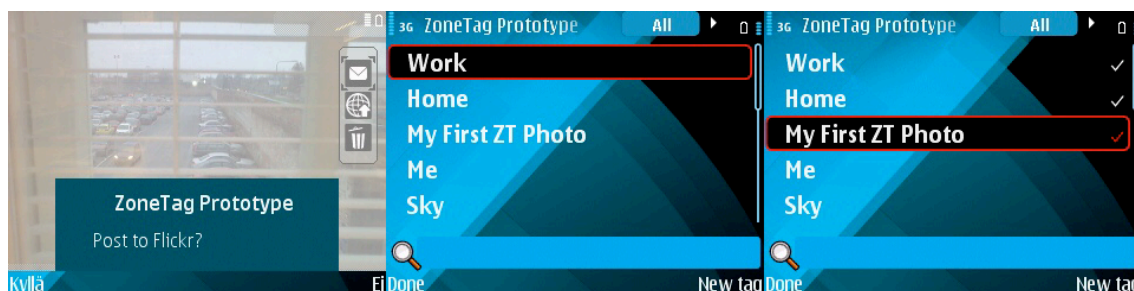
4.1.1. ZoneTag

ZoneTag on monipuolinen matkapuhelimeen tarkoitettu client –ohjelmisto. Se mahdollistaa kontekstitietoisien kuvien lataamisen matkapuhelimella verkkoon. ZoneTag mahdollistaa automaattisesti paikkatiedon lisäämisen kuviin, joita matkapuhelimella otetaan. ZoneTag tarjoaa myös automaattisia tagi ehdotuksia käyttäjälle. Ehdotukset haetaan verkosta eri lähteistä. Lähteinä toimii mm. media julkaisuja, käyttäjän itse lisäämiä tageja, käyttäjän sosiaalisen ja julkisen verkon tageja sekä paikkatietojen mukaan lähellä sijaitsevia rakennuksia, ravintoloita tai tapahtumia. Helppokäyttöisen rajapinnan ansiosta tagien lisääminen kuviin on helppoa ja näin voidaan muodostaa paremmat mahdollisuudet kuvien hakuun myöhemmin. [75; 76]

Valokuvaaminen on lisääntynyt lähivuosina paljon. Tähän syynä on kuvien ottamisen helppous ja kuvaamisen mahdollistavien laitteiden lisääntyminen (digitaalikamerat ja älypuhelimet) sekä muistin hintojen aleneminen. Kuitenkin metatiedon liittäminen kuvaan kätevästi kontekstitietona on yhä saavuttamatta. Tällaista ominaisuutta kuitenkin tarvitaan kuvien erottelemiseksi jälkikäteen kuvien paljoudesta johtuen. Täysin automaattista työkalua tarkan ja riittävän semanttisen metatiedon tuottamiseen ei ole vielä kehitetty. Manuaalisesti toimiva rajapinta metatiedon syöttöön on taas liian aikaa vievää erityisesti matkapuhelimilla. Tästä johtuen puoliautomaattisesti toimivia rajapintoja on yritetty kehittää yhä pidemmälle, etenkin tiedon tutkimisen ja kontekstin lisäämisen osalta. Puoliautomaattinen rajapinta tarkoittaa, että kuviin voidaan liittää jotakin kontekstia automaattisesti ja muuten

ehdottaa mahdollisia konteksteja, joita esimerkiksi kuvaan voidaan lisätä. [75; 76]

ZoneTag on sovellus automaattisen metadatan käsittelyyn ja sen avulla tuotettuihin tagiehdotuksiin. ZoneTag tukee kontekstiin perustuvaa metatietojen jakamista. Tämä toimii siten, että käyttäjä voi asettaa omat metatietonsa julkisiksi, minkä jälkeen ne ovat jaossa muiden käyttäjien kanssa. Tämän avulla saadaan ehdotettua muille käyttäjille toisten käyttäjien syöttämiä tageja, ja näin tagikirjastosta tulee ”rikkaampi”. Haastavuutena metatietosyötteiden suunnittelussa on tehdä siitä helppokäyttöinen sekä käyttäjille että kehittäjille. [75; 76]



Kuva 4.1. ZoneTagin toiminta ja taggaus.

ZoneTag mahdollistaa kuvien oton ja lataamisen helposti Flickr-verkkopalveluun (katso kuva 4.1.). Tähän ei vaadita parhaimmassa tapauksessa, kuin kaksi napin painallusta. Käyttäjä hyötyy palvelusta paljon, koska kuvaan tulee automaattisesti paikkatiedot. Lisäksi sovelluksessa on mahdollista lisätä tagit ZoneTagin omasta kirjastosta hakutoimintoa hyväksikäyttämällä. Tagit voidaan myös lisätä jälkikäteen myös Internetin kautta, joten niitä ei ole pakko syöttää kuvien latauksen yhteydessä. [75; 76]

ZoneTagin pääkomponentit ovat client -ohjelmisto, joka toimii Nokian S60-sarjan kamerapuhelimeissa ja Motorolan kamerapuhelimeissa, sekä ZoneTagin oma PHP/MySQL-serveri, joka tarjoaa kuviin paikkatiedon, ylläpitää jokaisen käyttäjän omaa tagikirjastoa, sekä prosessoi ladatut kuvat metatietoineen ja lähettää ne edelleen Flickr-palveluun. [75; 76]

ZoneTag-client toimii matkapuhelimesta tausta-ajona. Se valvoo, mihin tukiasemaan matkapuhelin on kytkettynä ja tarjoaa jatkuvasti paikkatietoa. Lisäksi se voi tarkkailla GPS -tietoja, joista saadaan edelleen tarkempaa paikkatietoa. Kun puhelin vaihtaa matkapuhelin tukiasemaa, client ottaa yhteyden serverille ja päivittää uudet tagiehdotukset uuden kontekstitiedon mukaisiksi. Muulloin päivitys tapahtuu 10 minuutin välein. Tagiehdotukset haetaan etukäteen puhelimeen, jottei verkko ruuhkaannu liiaksi latausvaiheessa ja toiminta tässä kohtaa toimii muutenkin nopeammin. Kun kuva on otettu, ZoneTag-sovellus tulee etualalle puhelimen näytössä ja tarjoaa näin ollen integroidun rajapinnan taggaamiselle, mikä mahdollistaa käyttäjälle nopean kuvan nimeämisen ja lataamisen. ZoneTag lähettää tagien ja kuvan lisäksi myös paikkatiedot serverille. Kun serveri saa uuden kuvan, muuntaa se clientin lähettämät paikkatiedot luettavaan muotoon. Tällöin myös Flickr-palvelu pystyy käyttämään niitä

oikein tageina. Mikäli ZoneTag ei osaa yhdistää matkapuhelin tukiaseman antamia tietoja kartalle, voi käyttäjä opastaa serveriä Internet-rajapinnan kautta, jolloin ZoneTag tallentaa opastetun paikan omaan serveriinsä. Uudet paikat lisätään myös käyttäjän sosiaaliseen verkkoon Flickr-palveluun. Paikan avulla saadaan myös käyttäjälle tärkeitä tageja paikkatietoon liittyen. Tagiehdotuksia tulee myös sosiaalisesta verkosta käyttäjien ja ystävien kautta sekä muilta käyttäjiltä, joilla on samanlaisia mieltymyksiä. Nämä ehdotukset tulevat ZoneTagin toimesta, eikä käyttäjä vaikuta siihen mitenkään. Kuitenkin käyttäjän kannalta tärkeimmät tagit ovat paikkatietojen ja ystävien kautta tulevat tagit, sekä omat aikaisemmin syötetyt tagit. Clientin lisäksi ZoneTag tarjoaa rajapinnan myös Flickr:n kanssa. Tämän ansiosta myös siellä voidaan tarjota samoja tageja, kuin clientissä. ZoneTag-serveri indeksoi ajoittain kuvia, jotka on lähetetty ZoneTag-clientilla. Jos kuviin on muutettu tai lisätty tageja, päivitetään ne käyttäjän omaan tagikirjastoon. [75; 76]

ZoneTag on ollut käytössä noin kolme vuotta, joten aivan uudesta sovelluksesta ei voida puhua. ZoneTagin olemassaolon aikana sovellusta on käyttänyt yli 1500 ihmistä ja sovelluksella on ladattu yli 100 000 kuvaa 60:ssä eri maassa. Noin 47,2 prosentissa kuvista oli ainakin yksi käyttäjän itse lisäämä tagisana, joka on kuvaan sopiva. 91 prosentissa tagatuista kuvista oli ainakin yksi tagisana, jota ei oltu syötetty kuvan lataamisen yhteydessä, eli tagi oltiin syötetty jo aikaisemmin tai tagi löytyi ZoneTagin omalta serveriltä tulleista tagiehdotuksista. Tämä puoltaa sekä kirjaston tehokkuutta, että omien metatietojen tallentamista kuvista toiseen. Tutkimuksessa kävi myös ilmi, että tageja ei juurikaan enää lisätä Flickr-palvelussa, jos ne on lisätty käyttäen ZoneTagia. Tämä kertoo puolestaan sitä, että kirjastosta saadaan riittävät tagit ja niitä ei enää jälkikäteen tarvitse muuttaa. [75; 76]

ZoneTagin tagikirjasto päivitetään puhelimesta manuaalisesti aina käyttäjän näin halutessa. Kirjasto sisältää seuraavanlaiset kategoriat:

- Kaikki: Kattegoria sisältää tagit kaikista kategorioista. Se sisältää siis kaikki tagit, jotka käyttäjä on koskaan lisännyt puhelimesta tai ZoneTagilla ladattuihin kuviin Flickr-palvelusta.
- Paikallinen: Kattegoriaan kuuluu käyttäjän tai sosiaalisen verkon syöttämät tagit, joita todennäköisesti käytetään nykyisessä paikassa.
- Viimeisimmät: Kattegoriaan kuuluu tagit, joita käyttäjä on lisännyt viimeisimmän 24 tunnin aikana.
- Tapahtumapaikka: Kattegoriassa on lähellä sijaitsevat paikat, joissa on tapahtumia. Tapahtumapaikat haetaan ”upcoming.org” –sivustolta.
- Tapahtumat: Kattegoria sisältää tapahtumat lähistöllä. Tapahtumat haetaan ”upcoming.org” –sivustolta, kuten edellä.
- Syötävä: Kattegoriassa ovat tageina lähellä sijaitsevat ravintolat.
- Termit: Kattegorian tagit generoidaan lähellä olevista tapahtumista erottelemalla tapahtumista erinäisiä termejä.
- Lisäksi käyttäjä voi itse muunnella omia tagi kategorioitaan RSS:n avulla. [75; 76]

4.2. Case: MoViE –projekti Pori Jazzeilla

Projektiin kuului Pori Jazzien yhteydessä suoritettavia pienempiä caseja. Nämä liittyivät lähinnä palvelimemme toimivuuden testaamiseen ja MoViE –clientin testaamiseen. Lisäksi selvitettiin yleisesti ihmisten innostusta MoViE:n kaltaisten sovellusten käyttöön. Casessa oli yhteistyössä erityisesti Pori Jazz, Tampereen teknillinen yliopisto, Turun yliopisto sekä Taideteollinen korkeakoulu, Porista. Tampereen teknillinen yliopisto vastasi pääosin teknillisestä toteutuksesta. TTY:n lisäksi casesta tuli arvokasta tutkimusmateriaalia myös yhteistyökumppaneille. Valmistelut aloitettiin hyvissä ajoin, noin 3 kuukautta ennen Jazz-festivaalien alkua. Caset itsessään kesti kaksi viikkoa. Pääosissa oli kaksi casea. Case I: Pori Jazzien kilpailu, jossa kuvattiin lyhyt video mobiililaitteella ja lähetettiin se Internet-sivustolle, sekä Case II: MoViE –client testaus. Jälkimmäisen casen testausryhmä koostui Jazzien henkilökunnasta.

Kilpailuprojektin ohessa Jazzeilla oli myös kolmen hengen ryhmä, joka keräsi kyselyitä ihmisiltä heidän matkapuhelimen käytöstään ja tarpeistaan. Kyselyiden tuloksista muodostettiin myös diplomityö. Kyselyyn vastasi yhteensä 162 henkilöä ja vastaajat olivat pääasiassa 21-30 vuotiaita. Kyselyssä pyrittiin kartuttamaan käyttäjien asenteita uutta teknologiaa kohtaan ja kuinka käyttäjät tulevat teknologiaa käyttämään. Huomion arvoista oli mitä ominaisuuksia käyttäjät käyttävät matkapuhelimellaan. Yli 50 prosenttia kyselyyn osallistuneista ovat käyttäneet matkapuhelimellaan kameraa, soittimia ja www-selainta. Näiden tulosten perusteella ihmiset ovat suuntautumassa enemmän mobiileihin palveluihin. Myös sosiaalisen median palvelut olivat saaneet suuren suosion (esimerkiksi Facebookia on käyttänyt yli 70 prosenttia vastanneista, joskin kyseinen tulos ei ole enää liitoksissa matkapuhelimeen). [42]

4.2.1. Case I: Kilpailu Pori Jazzeilla

Kilpailu oli Pori Jazzien ”pääkilpailu” vuonna 2009. Tästä johtuen pidimme monia palavereja Jazzien henkilökunnan kanssa, lähinnä ulkoasusta ja kilpailusivuston toimivuudesta sekä helppoudesta. Palavereita pidettiin noin 1-2 kuukaudessa. Palavereihin osallistui Jazzien myyntijohtaja Juha Miikkulainen ja myyntipäällikkö Niina Eeva.

Kilpailun tarkoitus oli ikuistaa kännykällä Pori Jazz 2009 – festivaalin hauskin, taiteellisin, kaihomielisin, romanttisin, ohikiitävin, iloisin tai koskettavin kohtaaminen. Lisäksi annoimme vapautta videoiden kuvaamiseen, koska videon ei tarvinnut olla festivaalialueelta tai Jazz-kadulta.



Kuva 4.2. SaxoPhone kilpailusivusto Pori Jazzeilla. Kuva videot-välilehdestä.

Kilpailun nimeksi tuli saxoPHONE, koska se kuvasti Jazzeja ja nimeen sai myös upotettua helposti mobiililaitteen ("phone"). Kilpailun Internet-sivut tehtiin mukailemaan Jazzien Internet-sivuja mahdollisimman paljon (katso kuva 4.2.). Saimme Pori Jazzin sivuille myös bannerin, josta pääsi suoraan kilpailun sivuille. Näin ihmiset löytäisivät helposti kilpailusivuille. Pori Jazz myös neuvotteli kilpailulle tilan, jossa oli päivystys arkipäivisin ongelmatilanteiden sattuessaa tai muuten apua tarvitseville. Kilpailun sivustot toteutettiin käyttäen jo olemassa olevia MoViE – sivustoja pohjana, jotka on luotu client -ohjelmistolle, lisäksi vain grafiikkaa ja muutamia lisätoimintoja. Kilpailusivustojen aloitussivulle lisättiin mm. Floobs-soitin, jonka ansiosta kaikki kilpailuun osallistuneet videot saatiin näkyville kätevästi. Tämän mahdollisti Floobsin

helppo käyttöliittymä ja videoiden lisääminen sivustolle. Tämän jälkeen sivuston luoma generoitu koodi vain lisättiin kilpailusivustolle. Kilpailun sivustolla oleva videotvälilehti pyrittiin tekemään mahdollisimman paljon YouTuben kaltaiseksi, jotta käyttäjien olisi helpompaa selata sivustoa. Tämä perustui YouTuben suosioon. Miltei kaikki ovat jossain tilanteessa käyttäneet kyseistä sivustoa.

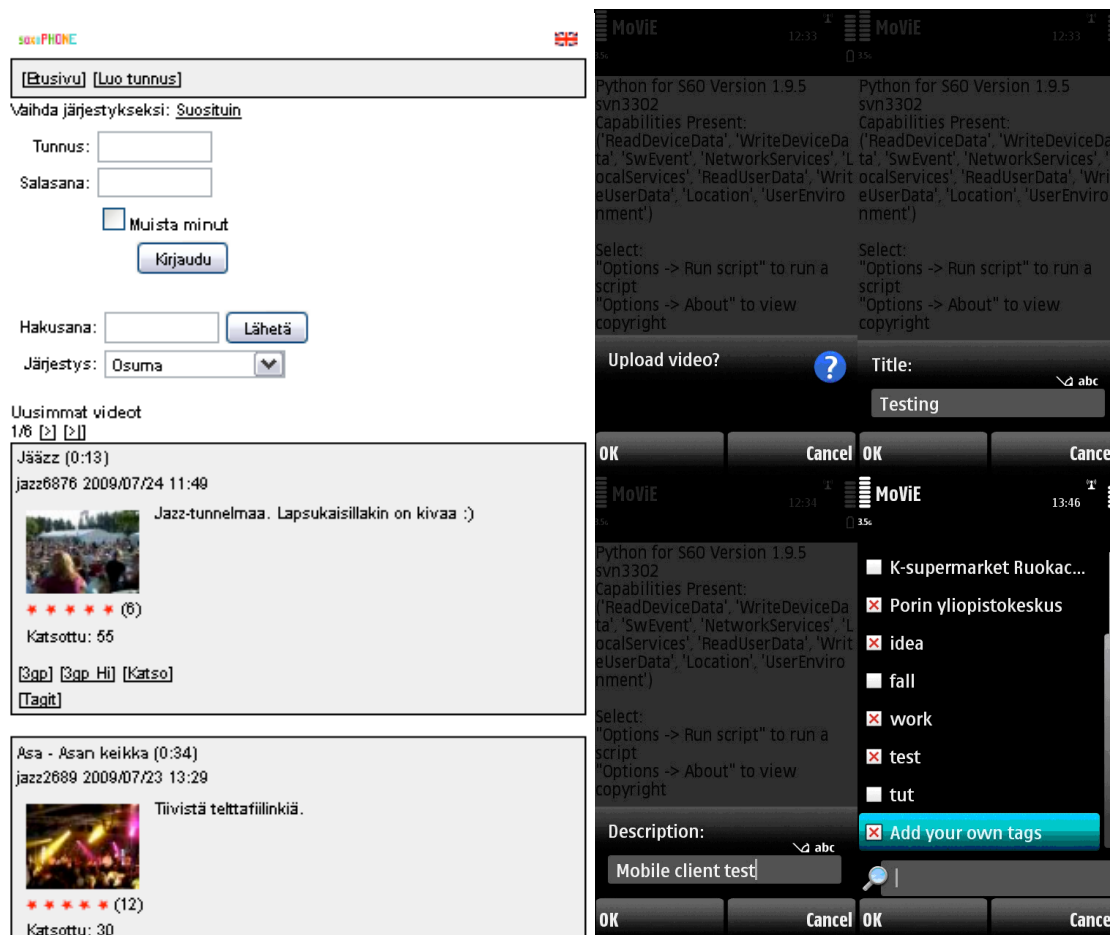
Kilpailu itsessään vaati rekisteröitymisen matkapuhelimella, jolloin saatiin mobiilius mukaan. Rekisteröitymiseen tuli vastaus tekstiviestillä, jossa oli käyttäjätunnus ja salasana. Näiden avulla pääsi kirjautumaan kilpailuun ja lataamaan videon. Myös palkinnot tulivat Jazzeilta, kolme kahdenhengen lippupakettia Pori Jazzin konsertteihin vuodelle 2010.

Kilpailun tavoitteena oli saada noin 100 videota kilpailuun, mutta tavoitteesta jäätiin hieman. Tämä saattaa johtua siitä, että ihmiset kokevat vaivalloiseksi kuvata videoita kännykällä ja sen jälkeen lähettää ne Internetiin erikseen. Tämä taas kannustaa MoViE –client ohjelmistoa, koska tässä videoiden lataus Internetiin tapahtuu automaattisesti, eikä käyttäjän näin ollen tarvitse tehdä sitä enää erikseen. Kyselylomakkeita analysoidessa löytyi myös eräs hieman yllättävä asia. Ihmisillä ei sen mukaan ole niin laajasti puhelimia, joilla voisi käyttää Internetiä. Tämä saattaa kertoa myös siitä, että ihmiset eivät ole halukkaita käyttämään Internetiä puhelimella tai siitä, että heillä on niin paljon ominaisuuksia puhelimessa, etteivät enää hallitse niitä kaikkia. Tämä puolestaan korostaa taas käyttöliittymän helppouden merkitystä.

Kilpailua markkinoitiin melko paljon, mahdollisimman monia maksuttomia kanavia hyväksikäyttäen. Juha Köönikkä ja Arttu Perttula kävivät mm. Jazz-radion haastattelussa ja tekivät jutun Porin Sanomiin sekä Satakunnan kansaan. Festivaaleilla jaettiin flyereitä, joissa oli lyhyt kuvaus kisasta sekä linkki kilpailun sivuille. Flyeristä tehtiin lisäksi mahdollisimman huomiota ja kiinnostusta herättävän näköinen.

4.2.2. Case II: MoViE –client testaus

MoViE –client testiryhmän tarkoituksena oli testata ohjelman toimivuutta, sekä kuvata samalla itselleen tärkeitä videoita Pori Jazzeilla. Testaukseen osallistui 8 henkilöä Pori Jazz festivaalin organisaatiosta. Client -ohjelmisto otettiin hyvin vastaan ja suurin osa oli innostuneita testaamaan ohjelmistoa. Testauksessa käytettiin Nokialta lainattuja matkapuhelimia. Matkapuhelimia oli 10 kappaletta ja ne olivat malliltaan N95 (5 kappaletta) ja Xpressmusic 5800 (5 kappaletta). Puhelimet asennettiin täysin toimintavalmiiksi ennen kuin ne luovutettiin testiryhmälle. Client -ohjelmistoa myös hieman yksinkertaistettiin, jotta sen kanssa toimiminen olisi entistäkin helpompaa (katso kuva 4.3.). Ohjelmisto valitsi tässä versiossa automaattisesti videoiden lataukseen käytettävän verkon sekä videon, joka palvelimelle ladataan. Hankimme puhelimiin myös liittymät, jotka kattaisivat datasiirron kahden viikon ajalta.



Kuva 4.3. MoViE sivusto ja client -ohjelmisto.

Käyttäjän ei tarvinnut tässä versiossa kuin käynnistää ohjelma päänäyttönäkymästä ja kuvata video. Tämän jälkeen syötettiin otsikko, kuvaus ja tagit. Näitäkään ei ollut pakko itse keksiä, koska palvelimeen oli rakennettu koodi, joka syötti automaattisesti viimeksi käytetyn otsikon ja kuvauksen sekä viimeisimmät käyttäjän tagit.

Client ei kuitenkaan toiminut Jazzien aikaan riittävän sulavasti, koska se käytti datansiirrossa 3G-verkkoa, EDGE- tai GPRS yhteyttä. Jazz festivaaleilla ihmisiä ja kännykän käyttäjiä on niin paljon, että puhelinverkko toimi liian hitaasti videoiden sulavaa lataamista varten. Tämän vuoksi client-testaajat eivät ladanneet yhtä paljon videoita, kuin olivat aluksi ajatelleet. Videoita tuli kuitenkin melko runsaasti, yhteensä 112 kappaletta. Videoista noin puolet ladattiin kännykän client -ohjelmiston avulla palvelimelle ja niihin syötettiin yhteensä 47 tagia. Suurin yhden henkilön kuvaama videoiden määrä oli 30 kappaletta.

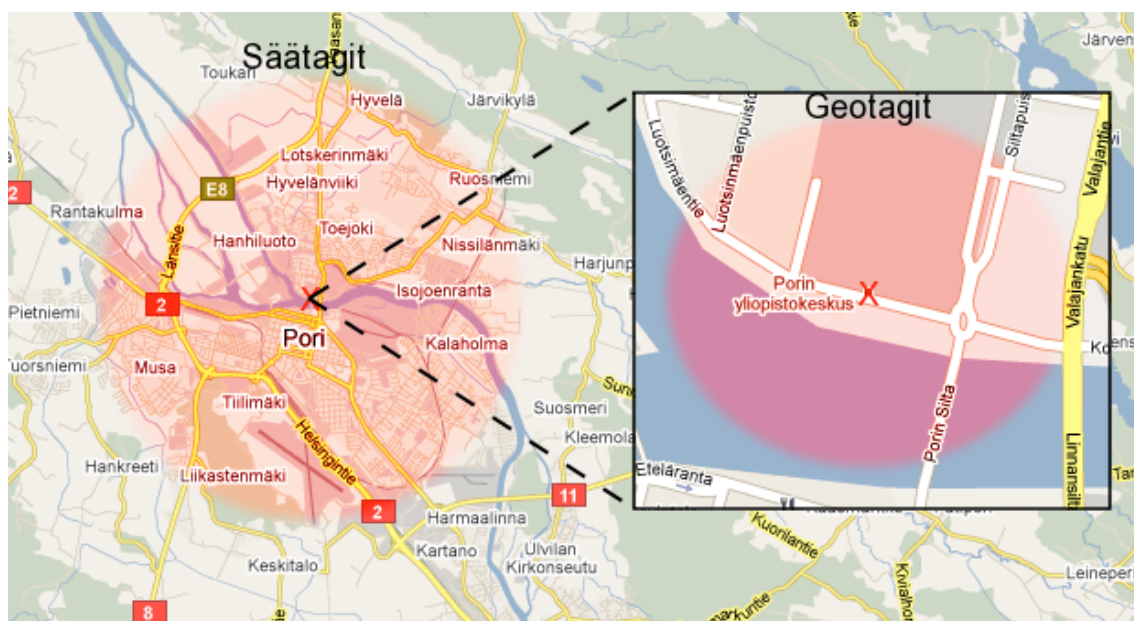
Festivaalien jälkeen pidimme vielä muutaman kokouksen sekä henkilökohtaisia haastatteluja osallistujien kanssa, missä käyttäjältä tuli palautetta kuvauksen työläisyydestä, vaikka alun opastustapahtumassa kaikki tuntuikin olevan erittäin helppoa ja selvää. Tämän arvelimme johtuvan siitä, kun monta videota päivässä kuvatessa voi alkaa tuntua työläältä kirjoittaa jokaiseen videoon erikseen otsikko,

kuvaus ja tagit, jos palvelimelta tulevat ehdotukset eivät soveltuneet käyttäjän tarkoitukseen. Ohjelmisto soveltuukin paremmin hieman satunnaisempaan käyttöön, kuin testiryhmillä oli. Lisäksi käyttökokemuksien jälkeen tyypillisesti vaatimukset kasvavat.

5. TAGGAUS MoViE – PROJEKTISSA

MoViE – projektissa on tarkoituksena kehittää matkapuhelimeen sovellus, jonka avulla voidaan kuvata video, lähettää se Internetiin palvelimelle ja liittää videoon lisäarvoa tuovaa kontekstietoa. Kontekstietoa videoon lisätään automaattisesti, puoliautomaattisesti ja manuaalisesti. Puoliautomaattinen lisääminen tässä tilanteessa tarkoittaa automaattisesti tulevien tagiehdotusten manuaalista hyväksymistä käyttäjältä. Videoita voidaan tämän jälkeen katsoa ja arvostella Internet-sivustolla. Palvelu on suunniteltu käytettäväksi ainoastaan matkapuhelimen kanssa, mutta sitä on mahdollista käyttää myös Internet-selaimella. Sivustosta on karsittu ylimääräiset efektit ja kuvat pois, jotta se olisi helppo ja kevyt matkapuhelimelle.

MoViE ohjelmisto ottaa tagitiedot käyttäjän viimeisimmistä tageista, xml-tiedostosta, paikkatiedon mukaan tietyltä alueelta sekä sään mukaan tietyltä alueelta. Alueet on määritelty sen mukaan, minkä kokoisia paikkoja voi olla. Tässä työssä paikan koon mittarina on käytetty keskikokoista kaupunkia (Pori), kuten kuvasta 5.1. nähdään. Tagit tulevat tietokannassa tehtävän laskutoimituksen mukaan, jotta ne olisivat mahdollisimman sopivia käyttäjälle.



Kuva 5.1. Paikkatiedon mukaan tulevat tagit kartalle hahmotettuna.

Säätiöiden mukaan tarjottavat tagit on valittu sen mukaan, miten tarkasti säätiötiedot yleensä annetaan. Pääasiassa säätiötiedot tulevat kaupungin mukaan, joten alue on rajoitettu keski-suuren kaupungin mukaan. Paikkatietojen mukaan tulevat tagit on rajoitettu sen mukaan, minkä kokoisia kaupungin alueita voi olla. Alue on rajattu noin (200 x 200)m kokoiseksi. Rajauksessa ongelmana oli määrittellä, mikä on riittävän kokoinen alue. Tapaissa voi ilmentyä virheitä myös sen mukaan, missä kohdin esimerkiksi sadepilvien reuna on tai esimerkiksi jos video on kuvattu aivan kaupunginosan reunassa. Nämä virheet ymmärretään ja todetaan, mutta niiden karsiminen ei olisi kannattavaa tagikirjaston koon pienenemisen kannalta, joten ne on jätetty huomioimatta.

Säätiöiden mukaan tulevissa tapaissa ei tule itse säätiötietoa, koska se tulee jo automaattisesti tietokantaan rakennetun koodin mukaan. Koodi käyttää säätiöiden hakemiseen GSM -solua. Mikäli tätä ei ole saatavilla, säätiötiedot haetaan GPS -tiedon mukaan. Säätiötiedot lisätään haun jälkeen kyseiseen videoon MoViE-sivustolla. Säätiöitä käytetään lähinnä sen vuoksi, että ihmiset voivat kuvata täysin samaa paikkaa aurinkoisella säällä ja sateisella säällä käyttäen erilaisia tapaista, sekä toisaalta ihmiset voivat kuvata täysin eri kohdetta samanlaisen sään vallitessa ja käyttäen täysin samoja tapaista. Säätiötiedon mukaan tulevat tagit ovat voimassa kaksi tuntia kuvaushetkestä eteenpäin, sillä sään oletetaan pysyvän kaksi tuntia samana. Tämäkään ei välttämättä ole totta. Jos taas säätiöiden mukaan tuleviin tapaisiin syöttää liian pienen ajan, voi se olla täysin turha kysely tietokannasta, koska välttämättä samassa kaupungissa ei tule montaa videota edes kahden tunnin sisään.

Xml-tiedostoa käytetään lähinnä erilaisten paikkojen lisäämiseen tagihetkentuksiin. Tietokantaan luetaan tietyin väliajoin xml-tiedosto, joka sisältää paikkojen nimiä sekä niiden koordinaatit. Paikkoja on koko Suomen alueelta noin 30 000 kpl. Xml-tiedosto on Internet-sivustolla ja sinne käyttäjät voivat lisätä itse paikkoja. Xml-tiedosto onkin melko kattava. Tiedosto sisältää suurimmaksi osaksi yritysten ja paikkojen nimiä. Tagitietojen keräämisestä ja niiden syöttämisestä kerrotaan lisää seuraavissa kappaleissa.

5.1. Taustatutkimukset

5.1.1. Tagikirjaston hyödyllisyys

Osana diplomityötä toteutettiin tutkimus, jossa ihmisille näytettiin kolme erilaista kuvaa. Kuvat valittiin siten, että ne eroaisivat toisistaan merkittävästi. Kuvissa oli paljon erilaisia mahdollisuuksia tagata eri asioita. Ihmisiltä kysyttiin mitä tagisanoja he käyttäisivät kuvaamaan näitä kuvia. Kysely toteutettiin sekä paperiversiona sekä Internet-kyselynä. Vastauksia kyselyyn tuli yhteensä 59 kappaletta. Näistä 16 vastasi paperiversioon ja 43 Internet-kyselyyn. Molemmissa kyselyissä oli samat kuvat ja ne

olevan hieman hankalampi kuin ensimmäinen ja kolmas kuva, jotta saataisiin hieman toisistaan eroavia tuloksia. Tässä kuvassa myös vastaukset erosivat toisistaan, mutta valtaosa kuitenkin käytti paljolti samoja tai samantyyllisiä sanoja.

Kolmas kuva oli taas valittu siten, että siinä oli yksi hallitseva näkymä, joka oli tori. Tätä sanaa olikin tagannut 86 prosenttia kaikista vastanneista. Toiseksi eniten kuvasta oli tagattu sanaa ”ihmiset”. Tätä oli tagannut 20 prosenttia kaikista vastaajista.

Tagikyselyssä oli mahdollista liittää jokaiseen kuvaan 5 tagisanaa, mutta monet vastaajat eivät kaikkia viittä sanaa käyttäneet. Monet olivat myös käyttäneet saman sanan eri taivutusmuotoja. Esimerkiksi viimeisessä kuvassa, jossa oli torinäkymä oli muutamat taganneet sanoja, kuten ”torinäkymä”, ”kauppatori” ja ”torilla”. Tämä on myös eräs ongelma automaattisten tagiehdotusten luomisessa, eli sanat pitäisi saada aina syötettyä perusmuodossa järjestelmään ja täten myös käyttäjälle.

Tuloksia analysoidessa ilmeni kuitenkin selkeästi tagikirjaston hyödyllisyys. Monet ihmiset käyttivät samantyyllisiä tai samoja sanoja tageja syötettäessä. Lisäksi monet olisivat saattaneet syöttää lisää tageja, jos olisivat saaneet ehdotuksina toisten syöttämiä tageja.

Tagien käyttö ei kuitenkaan vielä ole kovin yleisessä käytössä. Piritta Poikselän diplomityössä ”Mobiili sisällön luominen ja julkaisu osana sosiaalisen median palveluja” on kysymys tagien käytön yleisyydestä. Työssä tutkittiin vastaajien kiinnostusta tagien lisäystä kohtaan. Tähän kysymykseen vastasi 149 kyselyn täyttäjää. Kysymys kuului ”Oletko lisännyt tageja esim. kuviin, videoihin tai blogeihin?”. Kysymykseen vastanneista 49 prosenttia oli lisännyt tageja johonkin. Tämä vahvistaa päätelmää, että tagit eivät ole vielä kovin suosiossa, mutta toisaalta ihmiset saattaisivat käyttää tageja helpommin, jos he saisivat tagiehdotuksia. Näistä voisi löytyä sopivia tageja mediaan ja ne myös voisivat toimia esimerkkinä miten tageja käytetään. [42]

5.2. Tiedon kerääminen

Tutkimuksen tarkoituksena oli saada tagien syöttö mahdollisimman vaivattomaksi käyttäjän kannalta. Sovellus tarkkailee geotiedon avulla, missä video on kuvattu, mihin aikaan ja millainen sää kuvaushetkellä on ollut. Näiden tietojen mukaan pyritään saamaan mahdollisimman kattava tagilista käyttäjälle. Työssä pyrittiin tulokseen, jossa parhaassa tapauksessa käyttäjän ei tarvitse syöttää itse mitään tietoa ohjelmistoon. Tagit tulevat sekä itseltä, että muilta käyttäjiltä, joten varsinaista sisäänrakennettua kirjastoa ei palvelimelle ole luotu. Tagiehdotukset tulevat tietokantakyselyiden syötteenä puhelimeen.

Seuraavissa luvuissa keskitytään tiedonkeräämisentapoihin ja -tyyleihin palvelimella ja mobiililaitteissa.

5.2.1. Palvelin ja tietokannan tekninen toteutus

Palvelimella tarkoitetaan itse fyysistä tietokonetta, joka on verkossa. Palvelin tarkoittaa myös tietokoneohjelmaa, joka ottaa vastaan tietoja asiakasohjelmilta ja vastaa niiden lähettämiin tietopyyntöihin sovitulla tavalla. Palvelinta käyttävää sovellusta kutsutaan asiakkaaksi (client). Tässä tapauksessa asiakas on mobiililaitte, eli tarkemmin, matkapuhelin.

Palvelimelle toteutetussa tagikyselyissä käytetään tietokantakyselyitä ja niistä tulevia vastauksia tagien näyttämiseen. Tietokantakyselyt on toteutettu PHP:lla (HyperText Preprocessor), joka on varsin suosittu ja joustava ohjelmointikieli juuri tietokantapohjaisiin verkkosovelluksiin. PHP:ta käytetään lisäksi laajasti sen vuoksi, että se on niin sanottu *cross platform* –ohjelmisto. Tämä tarkoittaa sitä, että ohjelmisto toimii laajasti sekä Windows- että Unix-pohjaisissa käyttöjärjestelmissä. [77]

Tietoa kerätään PHP:lla tehtävien tietokantakyselyiden avulla käyttäjän jo tietämättään antamasta kontekstitiedosta. Kontekstitieto on tässä tapauksessa mm. käyttäjätunnus, paikkatiedot, säätiedot ja aika. Näiden tietojen avulla saadaan tietokannasta melko kattava henkilökohtainen tagikirjasto käyttäjälle. Tagikirjastosta tehdään pääsääntöisesti julkinen. Tämä sen vuoksi, että saataisiin kirjastosta mahdollisimman laaja.

Palvelimelle tallennetaan MoViE – palvelussa kaikki videoille ja käyttäjille merkittävä tieto. Palvelimella sijaitseva tietokanta sisältää 24 taulua ja niiden toiminta on ohjelmiston kannalta välttämätöntä. Tässä työssä on käytetty tauluja ainoastaan tagien hakuun. Tietokannan toteutus näkyy kuvassa 5.3.

uudelleen käytön yhdessä VIDEO_HAS_TAG -taulun kanssa. TAGI -taulu on välittömässä yhteydessä USERS- ja VIDEO_HAS_TAG -taulun kanssa. Nämä yhteydet ovat erittäin tärkeitä. Tässä työssä tietokannan TAGI -taulusta haetaan tageja erilaisten parametrien avulla. Tageja haetaan eri taulujen kautta, jotta saadaan määriteltä tarkemmin, mitä tagia juuri kyseisellä hakukerralla haetaan. TAGI -taulu on työn kannalta yksi tärkeimmistä tauluista.

VIDEO_HAS_TAG -taulu liittää käyttäjän lisäämät, sekä tietokannasta ehdotuksena tulleet tagit tiettyyn videoon. Tämä taulu mahdollistaa jo aiemmin lisätyn tagin liittämisen uuteen videoon, kuitenkin monistamatta tagia. Taulu sisältää kentät tagin id:lle, videon id:lle, käyttäjän id:lle sekä tagin lisäämisajalle. Näiden kenttien avulla voidaan tagien ja videoiden liitokset luoda. Tämän taulun avulla pystytään kaikki tagit liittämään yksiselitteisesti tiettyihin videoihin. Tämän taulun tärkeä piirre on myös tagin lisäämisaika videoon, koska sen avulla voidaan määritellä jo aiemmin lisätyn tagin (vertaa TAGI -taulun tag_added) liittämishetki videoon. Tätä käytetään säätietojen määrittelyyn, missä on aikarajana kaksi tuntia.

VIDEO -taulu sisältää tiedot videoista ja toimii samalla toisena kahdesta päätaulusta. Myös VIDEO -taulu ohjaa monien muiden taulujen toimintaa sekä toimii niin sanotusti välikätenä taulujen toiminnassa. Taulu sisältää myös tiedon ketä videon on palvelimelle lisännyt. Taulua käytetään tässä työssä muun muassa otsikko- ja kuvaehdotuskyselyiden tekemiseen. VIDEO -taulu on välittömässä yhteydessä GEOTAGI -taulun kanssa. Siksi VIDEO -taulua käytetään niin sanotusti välikätenä hakemaan esimerkiksi paikkatietoihin perustuvia tageja. Lisäksi taulu tekee videoista yksiselitteisesti jonkin käyttäjä palvelimelle lataamia.

GEOTAGI -taulu on luotu paikkatietojen antamista varten. Paikkatiedot haetaan GPS:n avulla tai jos tämä tieto ei ole saatavilla käytetään GSM -solua paikkatietojen määrittämiseen. Näiden tietojen avulla haetaan paikkatietoon perustuvia tageja. GEOTAGI -taulun tiedot välitetään taulujen välisten suhteiden mukaan myös TAGI -taululle ja tämä mahdollistaa tagien helpon haun.

GEO_PLACES -taulu on rakennettu irralleen muusta kokonaisuudesta, koska sitä päivitetään erillään muusta kokonaisuudesta. Taulu sisältää paikkojen nimiä sekä nähtävyyksiä. Tähän tauluun päivitettävä tiedosto haetaan geo-position.com -sivustolta. Sivusto tarjoaa paikkatiedot jo xml-muotoisena ja niitä on siksi helppo päivittää tietokantaan. Xml-tiedosto päivitetään tietyn aikavälein. Taulun tehtävä on luoda paikkatiedon mukaan mahdollisimman tarkkoja tagiehdotuksia käyttäjälle. Tämä on mahdollista toteuttaa erilleen muusta kokonaisuudesta, koska taulu sisältää jo itsessään paikkojen GPS -koordinaatit.

5.2.2. Mobiililaite

Sana mobiililaite (esimerkiksi matkapuhelin tai PDA, Personal Digital Assistant) voi kuulostaa melko vieraalta tai tuntemattomaltakin asialta kokemattomamman käyttäjän

mielestä. Mobiiliudella tarkoitetaan sitä, että tietoa voidaan vastaanottaa ja välittää paikasta riippumatta. Mobiililaitte on siis tietoliikenneyhteydessä langattomasti. Mobiililaitteilla voidaan olla yhteydessä tietoverkkoihin ja tätä kautta tietokoneisiin, puhelimiin tai muihin laitteisiin. [78]

Mobiililaitteelle tulevat tagit luetaan tietokannasta ennalta määrätystä paikasta ja ne tulevat luettelosyötteenä puhelimeen. Palvelimen tehtyä tietokantakyselyt se lähettää saadut tulokset palvelimella sijaitsevaan tiedostoon. Tiedosto muuntaa tulokset xml-muotoon, josta mobiililaitteen on ne helppo lukea. Mobiililaitteessa on käytetty ohjelmointikielenä Mobile Python –kieltä (Python for S60, PyS60), joka soveltuu erinomaisesti ”tee-se-itse” kokeiluihin, prototyyppien tekoon ja peleihin. Tässä tapauksessa Pythonia on käytetty demonstraatio sovelluksen tekoon. Python on myös hyvä valinta, koska se ei vaadi puhelimelta paljoa tehoa ja sitä on helppo ja nopea muokata. Tämä perustuu Python –koodin yksinkertaiseen syntaksiin, korkean tason tietorakenteisiin ja keveyteen. Python toimii Symbian käyttöjärjestelmässä. [79; 80]

5.3. Tiedon liittäminen ja siirtäminen

Kontekstitieto voi olla henkilökohtaista. Käyttäjä ei aina ole halukas luovuttamaan kaikkea tietoa, jota ohjelmisto on valmis ottamaan. Etenkään nykytekniikan mahdollistaessa kaiken tiedon jakamisen ja säilyttämisen. Kerran julkaistua tietoa on usein mahdotonta estää leviämistä. Kontekstitietoon liittyvä suuri ongelma onkin yksityisen tiedon ja henkilökohtaisen kontekstitiedon eriyttäminen toisistaan. Ratkaisuna tämän kaltaisiin ongelmiin on esimerkiksi kontekstilajikohtainen lupamenettely. Tämä rajaa tietoja, joihin ohjelmistot pääsevät käsiksi ja jotka ne näin ollen voisivat liittää kontekstiin. Käyttäjän tulisi voida luottaa siihen, ettei tietoa käytetä väärin tarkoituksiin. Kuitenkin on käytännössä mahdotonta olla täysin varma, ettei yksityisiä tietoja leviäisi ilman käyttäjän lupaa.

Diplomityössä toteutettu ohjelmisto on rakennettu ajatellen sitä, että kukaan ei syöttäisi palveluun mitään tietoa, jonka haluaisi olevan yksityistä. Kaikki tagisanat on luettavissa ohjelmiston Internet-sivustolta, johon kaikilla on vapaa pääsy. Tästä johtuen myös toisten tagisanoja voi ehdottaa muille ajasta ja paikasta riippuen.

Kerätty tieto liitetään PHP -koodiin, josta mobiililaitteeseen asennettu ohjelmisto käy sen lukemassa. Koska tietoa on paljon, joudutaan se liittämään koodiin erilaisien silmukoiden avulla. PHP -koodin tekniseen toteutukseen paneudutaan tarkemmin luvussa 5.4.1. Tekninen toteutus palvelin ja client.

Tieto tulee myös siirtää asiakkaan ja palvelimen välillä. Tähän tarkoitukseen voi käyttää, joko WLAN -yhteyttä (Wireless Local Area Network) tai mobiililaitteen liittymän omaa Internet-yhteyttä. Näistä suositeltavampaa olisi käyttää WLAN –yhteyttä, koska videotiedostot voivat olla melko suuren kokoisia ja WLANilla tiedonsiirtonopeus on huomattavasti suurempi kuin matkapuhelimen 3G -verkossa. Siirrot ovat turvallisia siksi, että jokaisessa siirrossa on käytössä henkilökohtainen

käyttäjätunnus ja salasana.

5.3.1. Tagipalvelu palvelimella

Palvelimella pitää varmistaa, ettei sama tieto tule moneen kertaan käyttäjälle. Ei olisi hyödyllistä, jos sama tagi ilmestyisi moneen kertaan käyttäjälle. Sama tagi on silloin ja vain silloin täysin sama, kun tagin id ja nimi on sama. Tämä tarkoittaa sitä, että samanniminen tagi voi tulla luetteloon. Tagi ei kuitenkaan ole sama, jos eri käyttäjät ovat syöttäneet saman sanan. Tähän on päädytty siksi, että käyttäjät voivat valita samoja sanoja tarkoittaen eri asiaa. Jos täysin sama sana (minkä id ja nimi olisi samoja) tulisi moneen kertaan tagiehdotukseksi, hidastaisi tämä tiedon tuontia mobiililaitteelle ja teettäisi lisäksi turhaa työtä. Tiedon monistumisongelma on ratkaistu vertailemalla jonoja, jotka tulevat tietokannasta. Tässä tapauksessa otetaan aina kerrallaan yksi jono käsittelyyn ja verrataan sitä seuraavaan ja poistetaan mahdolliset duplikaatit. Duplikaattien poistamisen jälkeen tieto syötetään yhtenä jonona tietokannassa sellaiseen paikkaan, josta mobiililaitte käy sitä lukemassa.

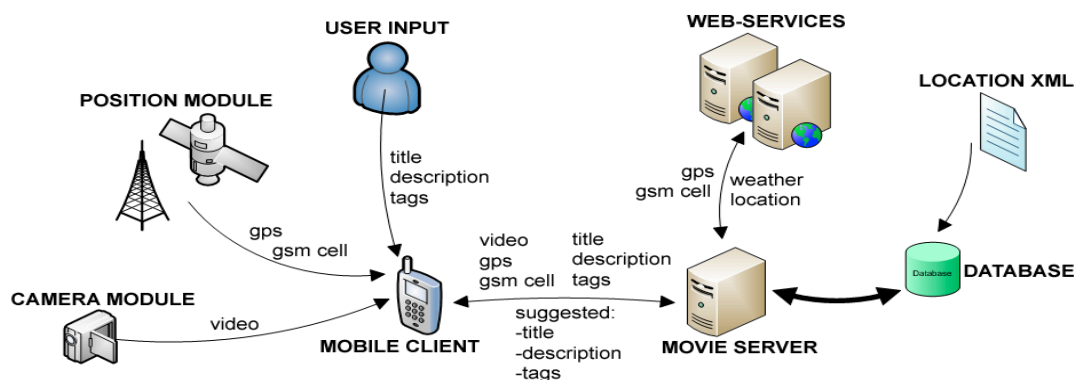
5.3.2. Tagipalvelu mobiililaitteessa

Mobiililaitteeseen on rakennettu ohjelma, joka käy lukemassa palvelimelta tietystä tiedostosta kaikki, mitä sinne on syötetty. Tagipalvelu ei siis varsinaisesti toimi mobiililaitteessa, vaan yhteistyössä palvelimen kanssa. Mobiililaitteen tehtävänä on tarjota palvelimelle kontekstitietoa, jonka avulla tagipalvelu toimii.

5.4. Tagi palvelun tekninen toteutus

Palvelin ehdottaa loogisesti erilaisia tageja erilaisten syötteiden johdattelemana. Suurin merkitys tagikirjaston tekemisessä on GPS -tiedoilla. MoViE-client testaajat vaikuttivat paikkatietoon perustuvien tagien tärkeyteen paljon, koska lähes kaikki korostivat juuri kyseistä asiaa tageissa. Käyttäjät olisivat halunneet tagata tarkkaan oman paikkansa kartalla ja käyttäneet tageina juuri sen paikan nimeä, missä videon olivat kuvanneet. Lisäksi kaikki olisivat toivoneet tagiehdotusten olevan niin hyviä, ettei käyttäjän itse tarvitsisi kirjoittaa tageja client -ohjelmistossa. Tagien kirjoittaminen matkapuhelimella koettiin hankalaksi ja aikaa vieväksi. Tämän vuoksi tässä työssä on panostettu paikan mukaan tulevien tagien mahdollisimman laajaan saatavuuteen.

MoViE-client käyttää hyväkseen serveriä sekä puhelimen sisään rakennettuja ominaisuuksia tagien hakemiseksi ja tarkentamiseksi. MoViE-client on tässä mielessä hyvinkin monipuolinen sovellus. Alla oleva kuva selittää tarkemmin koko MoViE:n toimintaa tagien ja kontekstitiedon osalta.



Kuva 5.4. MoViE:n toimintakaavio.

Kuvasta 5.4. ilmenee miten MoViE hakee eri paikoista tarvittavat kontekstit videoille. Konteksteja haetaan mm. Google- ja Yahoo!- palvelimelta, MoViE:n omasta tietokannasta sekä matkapuhelimen avulla GPS ja GSM -solutiedoista.

Aluksi käyttäjä käynnistää MoViE-client sovelluksen. Sovellus kysyy ensin käyttäjältä mitä verkkoa käytetään. Tämän client käynnistää mobiililaitteen oman kameran kuvaamista varten. Samalla kun kamera on käynnistetty, client hakee GPS paikkatiedot sekä GSM -solutiedot. Kun kuvaus on lopetettu, client tarkastaa, mikä on viimeisin video mobiililaitteessa ja lataa sen palvelimelle. Videon lataamisen yhteydessä lähetetään myös kontekstiatietoa videosta. Kontekstietieto sisältää seuraavat asiat:

- Käyttäjän ID (tunnus ja salasana)
- GPS tiedot
- GSM -solutiedot
- Kuvaus aika
- Videotiedosto

Kontekstietiedon lähettämisen jälkeen palvelin suorittaa kyselyitä tietokannasta clientilta saatujen tietojen perusteella. Kyselyiden tarkoitus on löytää mahdollisimman hyviä tagiehdotuksia käyttäjälle sekä ehdotus käyttäjän kuvaaman videon otsikoksi ja selitteeksi. Tämän jälkeen ehdotetut tagit, otsikko ja kuvaus lähetetään takaisin mobiililaitteelle. Mobiililaitteella käyttäjä valitsee videolle sopivimmat tagit sekä mahdollisesti korjaa otsikon ja kuvauksen mieleisekseen. Seuraavaksi käyttäjä lähettää tiedot takaisin palvelimelle. Kaikki videoon liittyvät tiedot tulevat näkyviin MoViE sivustolla. Sivustolla siis näkyvät kuvaajan tiedot, video, tagit, kuvauspaikka ja sää.

Toiminta kuvassa 5.4. Location XML tarkoittaa xml-tiedostoa, joka ladataan tietyin väliajoin Internetistä. Tiedosto sisältää paikkojen nimiä koordinaatteineen ja selityksineen. Tämän tiedoston avulla saadaan tagikirjastosta kattavampi.

5.4.1 Tekninen toteutus palvelin ja client

MoViE-palvelin koostui monesta tiedostosta, jotka ovat riippuvaisia toisista. Tässä työssä kuitenkin muokattiin pääosin vain yhtä tiedostoa, joka sisältää tietokantaan tehtävät kyselyt. Kyselyiden tulokset lähetettiin vielä eteenpäin toiseen tiedostoon, jonka avulla tiedot siirretään matkapuhelimeen. Seuraavassa osittain teknistä toteutusta miten tagikirjasto on rakennettu.

```
3909 function demoSuggestTags($usr_id, $vd_id, $geo_lat=0, $geo_lng=0)
```

Kuva 5.5. Palvelimen database.php tiedoston funktion määrittely

Aluksi luodaan funktio mitä käytetään tagikirjaston tekoon (kuvassa 5.5.). Funktion nimi on demoSuggestTags ja se käyttää sisälle rakennetuissa lausekkeissaan seuraavia arvoja:

- `$usr_id`, joka on käyttäjän id
- `$vd_id`, joka on videon id
- `$geo_lat`, joka on videon kuvauspaikan pituusasteen arvo
- `$geo_lng`, joka on videon kuvauspaikan leveysasteen arvo

Kyseisten arvojen avulla voidaan tehdä tietokantakyselyitä ja näin saadaan halutut arvot.

```
3910 //----- tagit -----
3911 global $mysqli;
3912 // SQL - kysely || DESC = käänteinen järjestys
3913 $Kysely = "SELECT tag_name, tag_id FROM tagi WHERE tag_owner=?
ORDER BY tag_added DESC LIMIT 5";
```

Kuva 5.6. Käyttäjän viimeisimpien tagien kysely tietokannasta.

Funktion määrittelyn jälkeen luodaan yhteys tietokantaan komennolla ”**global \$mysqli**”, ja tämän jälkeen voidaan tehdä ensimmäinen kysely tietokannasta. Ensimmäinen kysely tiedustele palvelimelle lähetettyjen videoiden joukosta tietyn käyttäjän (videon kuvaajan) 5 viimeisintä tagia sekä niiden id:t, eli yksilöllisen tunnusteen (kuva 5.6.).

Tämä kysely tehdään, koska monet käyttävät tageissa usein samoja tai samantyyllisiä tageja. Kyselyn ansiosta sellaiset henkilöt, jotka käyttävät samoja tageja, esimerkiksi kuvatessa monta videota samasta tapahtumasta, pystyvät syöttämään ne ohjelmaan nopeammin.

```

3914 //kyselyn määrittely
3915 if ($stmt = $mysqli->prepare($Kysely)) {
3916     $stmt->bind_param('i', $usr_id);
3917     $stmt->execute();
3918     $stmt->bind_result($tag_name, $tag_id);

```

Kuva 5.7. Palvelimen suorittaman tietokantakyselyn määrittäminen.

Kun tietokantaan tehtävä kysely on määritelty, voidaan kysely suorittaa. Kysely suoritetaan kuvasta 5.7. näkyvällä tavalla, koska se on standardina myös muissa kyselyissä. Kyselyn määrittelyt:

- `prepare`, valmistaa kyselyn
- `bind_param`, sitoo kyselyyn arvon, joka on määritelty sulkeissa
- `execute`, toteuttaa kyselyn edellä mainituilla arvoilla
- `bind_result`, sidotaan saadut tulokset näihin muuttujiin

```

3919 $i = 0;
3920 while ($stmt->fetch()) {
3921     $tag[$i]['tag_name'] = $tag_name;
3922     $tag[$i]['tag_id'] = $tag_id;
3923     $i++;
3924 }
3925 $stmt->close();
3926 } else {
3927     echo $mysqli->error;
3928 }

```

Kuva 5.8. Palvelimen tietokantakyselyn tulosten määrittely.

Useimmat tulokset sisältävät monia arvoja. Nämä arvot on laitettava jonoon (array), jotta niitä olisi helpompi käsitellä. Tässä suoritetaan tulosten jonoon laittaminen. Jonot nimetään halutulla tavalla, jotta niitä on helpompi käsitellä myöhemmissä vaiheissa. Koodissa olevat käskyt merkitsevät seuraavaa:

- `fetch`, hakee arvot niin kauan, kun määrittely on tosi
- `close`, sulkee kyselyn

Eli alussa määritellään muuttuja ”`$i`” arvoon nolla. Tämä on arvo, joka kasvaa jokaisella hakukierroksella kuvasta 5.8. näkyvällä määrätyllä tavalla. Aina yhdellä hakukierroksella haetaan tuloksista yksi arvo (”`$tag_name`”) ja syötetään se määriteltyyn jonoon (”`$tag`”). Muuttuja ”`$i`” määrittää paikan, mihin kohtaan jonoon kyseinen arvo laitetaan. Kyseinen toiminta tehdään, koska molemmat tiedot halutaan pitää jonon samassa alkiossa. Tämä taas johtuu siitä, että samalla kierroksella haetut tiedot liittyvät merkittävästi toisiinsa. Koska jonoon syötetään aina kaksi arvoa, on ne erotettava toisistaan jollain tavalla. Arvot erotetaan attribuuteilla, jotka tässä kyselyssä ovat ”`tag_name`” ja ”`tag_id`”. Jos alussa määritelty kysely ei jostain syystä onnistu, palautetaan tulosten sijaan virheilmoitus.

```

3929 //----- title & description -----
3930 $Kysely = "SELECT vd_name, vd_description FROM video WHERE
vd_usr_id=? AND vd_state=3 ORDER BY vd_id DESC LIMIT 1";
3931 if ($stmt = $mysqli->prepare($Kysely)) {
3932     $stmt->bind_param('i', $usr_id);
3933     $stmt->execute();
3934     $stmt->bind_result($vd_name, $vd_description);
3935     $stmt->fetch();
3936     $stmt->close();
3937 } else {
3938     echo $mysqli->error;
3939 }

```

Kuva 5.9. Palvelimen tietokantakysely, otsikon ja kuvauksen haku.

Kuvassa 5.9. haetaan tietokannasta videolle otsikko ja kuvaus ehdotukset. Toteutus tehdään, jos käyttäjä haluaa käyttää samaa tai melkein samaa otsikkoa ja kuvausta videoille. Tämä on käytännöllinen varsinkin, jos käyttäjä kuvaa monta videota lyhyen ajan sisään ja niiden aiheet ovat samantyyllisiä.

Kysely valitsee tietokannasta ainoastaan yhden videon ("LIMIT 1"), joka on käyttäjän kuvaama ("WHERE vd_usr_id=?"). Samalla määrätään myös, että video on käyttäjän uusin video ("ORDER BY vd_id") ja sen lataaminen palvelimelle on onnistunut ("vd_state=3"). Kun kyselyssä kysytään ainoastaan yhtä arvoa, ei tarvita silmukkaa, joka järjestäisi arvot jonoon.

Tästä kyselystä tulee siis palautteena ainoastaan käyttäjän palvelimelle lataaman viimeisimmän videon otsikko ja kuvaus. Tieto tulee automaattisesti syötteenä client-ohjelmistoon kuten kuvasta 5.30. voi havaita.

```

3940 //----- geotagi (1) -----
3941 $muuttuja_lng = 0.0025;
3942 $muuttuja_lat = 0.005;

```

Kuva 5.10. Tietokantakyselyn muuttujien määrittely.

Kuvassa 5.10. valmistellaan geotageja varten muuttujat. Muuttujat on valittu sen perusteella, että ne muodostavat tietyn kokoisen alueen. Alue ei ole täysin ympyrä. Tämä on todettu ja ymmärretty. Ympyrän muodostavan alueen luominen ei ole tässä työssä pääasia, vaan alueen järkevä rajaaminen tietyn kokoiseksi. Alue on siis erimuotoinen siirryttäessä lähemmäs päiväntasaajaa. Tämä johtuu pituus- ja leveysasteen muodosta.

```

3944 $muuttuja1 = $geo_lat-$muuttuja_lat;
3945 $muuttuja2 = $geo_lat+$muuttuja_lat;
3946 $muuttuja3 = $geo_lng-$muuttuja_lng;
3947 $muuttuja4 = $geo_lng+$muuttuja_lng;

```

Kuva 5.11. Tietokantakyselyn muuttujien laskutoimitukset.

Kuvassa 5.11. muokataan videon kuvauspaikan koordinaatteja siten, että alueesta saadaan hieman suurempi seuraavaa tietokantakyselyä varten. Alueen koko on muutoksen jälkeen halkaisijaltaan noin 200 metriä.

```

3949     $Kysely = "SELECT geo_vd_id FROM geotagi WHERE geo_lat BETWEEN
? AND ? && geo_lng BETWEEN ? AND ? ORDER BY geo_vd_id DESC LIMIT 6";
3950     if ($stmt = $mysqli->prepare($Kysely)) {
3951         $stmt->bind_param('ssss', $muuttuja1, $muuttuja2,
$muuttuja3, $muuttuja4);

```

Kuva 5.12. Palvelimen tietokanta kysely paikkakohtaisiin tageihin

Kuvassa 5.12. toteutetaan ensimmäinen osa kaksiosaisesta kyselystä, jonka tarkoituksena on löytää tiettyssä paikassa syötetyt tagit. Muuttujina kyselyssä käytetään aiemmin määriteltyjä arvoja.

Kyselyssä haetaan aluksi tiettyssä paikassa ladattujen videoiden id:t geotagi taulusta. Määrittelyn mukaan haetaan kaikkien videoiden id:t, jotka ovat tiettyjen arvojen sisällä. Tietojen hakeminen aloitetaan uusimmasta arvosta. Kyselyyn on liitetty rajaus, jos videoita on kuvattu samassa paikassa useita, hakee kysely ainoastaan kuusi viimeisintä arvoa. Muuttujat syötetään kyselyyn string muotoisina, koska samat arvot ovat myös tietokannassa string muotoisina. Tässä siis noudatetaan samaa formaattia.

Kuten edellisissäkin kohdissa tuloksena saadut arvot syötetään jonoon, jota käytetään seuraavassa kyselyssä hyväksi. Tämän kyselyn tulokset tallennetaan "\$numerot"-nimiseen jonoon.

```

3963     //----- geotagi (2) -----
3964     $j = 0;
3965     for ($i = 0; $i < count($numerot); $i++) {

```

Kuva 5.13. Palvelimen tietokantahaku, geotagien haku toinen osa, alustusvaihe.

Koska edellisen kyselyn tuloksia on useita ja ne ovat jo jonossa, täytyy seuraavan kysely toteuttaa silmukassa. Aluksi alustetaan muuttujat "\$j" ja "\$i". Ensimmäistä muuttujaa käytetään myöhemmin silmukan sisällä olevassa toisessa silmukassa, mutta muuttuja "\$i" käytetään "\$numerot" -jonon tarkastelussa. Silmukka toistaa seuraavaa kyselyä siis niin kauan kunnes koko "\$numerot" -jono on käsitelty läpi. Käsittely alkaa jonon ensimmäisestä alkioista.

Silmukan ehdot kuvassa 5.13. rivillä 3965 siis ensin alustavat muuttujan. Tämän jälkeen lasketaan montako alkioita jonossa "\$numerot" on ja lopussa siirrytään aina seuraavaan alkioon, kun edellinen alkio on käsitelty läpi.

```

3966     $Kysely = "SELECT DISTINCT tagi.tag_id, tagi.tag_name FROM
video_has_tag, tagi WHERE video_has_tag.vtag_tag_id = tagi.tag_id AND
video_has_tag.vtag_vd_id = ?";

```

Kuva 5.14. Palvelin tietokantakysely, geotagit toinen osa.

Geotagien kyselyn toinen osa käsittelee kyselyssä tietokannan kahta eri taulua ”video_has_tag” ja ”tagi”. Tämä kysely olisi voinut tehdä myös kahdessa eri osassa, mutta tilan säästämiseksi ja kyselyn yksinkertaisuudesta johtuen on kysely toteutettu yhdessä osassa (kuva 5.14.).

Kyselyn ehtona on, että tauluissa ”video_has_tag” ja ”tagi” on tageilla oltava samat id:t. Taulujen nimet ovat hieman erilaiset havainnoin vuoksi, mutta tarkoittavat kuitenkin samaa asiaa. Tageille tulee id silloin, kun se syötetään ensimmäistä kertaa tietokantaa ja sama id seuraa sille kuuluvaa tagia tietokannan taulusta toiseen. Toinen ehto kyselyssä on, että ”video_has_tag” taulussa ”vtag_vd_id” on oltava sama, kuin kyselyn ensimmäisen osan tulos.

Kysely hakee tietokannan taulusta ”tagi” tuloksensa. Tuloksena otetaan tagin id ja nimi. Lausekkeessa käskyllä ”DISTINCT” tarkoitetaan, että tuloksiksi hyväksytään ainoastaan yksi samalla id:llä ja nimellä oleva tulos. Tämä poistaa duplikaatit. Tuloksena siis tulee tagin nimi ja id jotka ovat uniikkeja arvoja. Tulokset syötetään saamaan jonoon ja ne erotetaan attribuutin avulla, kuten edelläkin.

Seuraavassa kyselyssä haetaan säähän liittyviä tageja. Kysely aloitetaan samoin kuten geotagien kysely, eli määrittämällä muuttujat miltä alueelta tageja haetaan (kuva 5.15.). Säätagien kyselyssä aluetta on suurennettu hieman, johtuen siitä että sää on suuremmalla paikalla sama. Säätagien kysely on toteutettu keskisuuren kaupungin kokoiselle alueelle (vertaa Pori).

```
3993 | $Kysely = "SELECT geo_vd_id FROM geotagi WHERE geo_lat BETWEEN  
? AND ? && geo_lng BETWEEN ? AND ? ORDER BY geo_vd_id DESC LIMIT 6";
```

Kuva 5.15. Palvelimen tietokantakysely, säätagien kyselyn ensimmäinen osa.

Säätagien tietokantakyselyn ensimmäisessä osassa haetaan taas ”geotagi” – taulusta videon id:t. Arvot jotka on määriteltä aikaisemmin, syötetään kyselyn ehdoksi. Tässä kohdassa tarkistetaan siis missä kohdassa video on kuvattu ja sen ympäriltä lasketaan tietty maantieteellinen alue jonka ympärillä syötetyt tagit otetaan mukaan tuloksiin. Kyselyyn on syötetty tuloksia rajaava määre ”LIMIT 6”. Tuloksien määrää rajaava käsky on syötetty, ettei tagikirjasto kasva liian suureksi. Tämän kyselyn myöhemmässä vaiheessa asetetaan vielä aikarajaus, joka rajaa tuloksia edelleen osuvammaksi käyttäjälle.

Tuloksena siis tulee 6 videon id:t, jotka ovat ”geotagi” – taulussa. Nämä tulokset syötetään taas jonoon myöhempää kyselyä varten.

```
4009 | $x = time();  
4010 | $y = $x - 7200;
```

Kuva 5.16. Palvelin, muuttujien määrittely tulevaan kyselyyn.

Tässä (kuva 5.16.) määritellään seuraavan kyselyn aikaväli. Muuttujaan ”\$x” otetaan aika, milloin kyselyssä käsiteltävä video on kuvattu ja lähetetty palvelimelle. Se

saadaan helposti funktiolla "time()". Tämän jälkeen vähennetään ajasta 7200 sekuntia. Arvo on sekunteina, koska tietokanta tallentaa ajan automaattisesti sekunteina. Tämä ajanlasku on lähtenyt 1.1.1970 keskiyöllä ja se on Greenwich aikaa. Eli 7200 sekuntia on sama asia kuin 2 tuntia. Tämä aika on ajateltu olevan realistinen aika sille, että sää pysyy suhteellisen samana. Aika on myös sellainen, että sinä aikana käyttäjät saattavat kuvata samalla alueella useampia videoita. [81]

```
4013      $Kysely = "SELECT DISTINCT tagi.tag_id, tagi.tag_name FROM
video_has_tag, tagi WHERE video_has_tag.vtag_tag_id = tagi.tag_id AND
video_has_tag.vtag_vd_id = ? AND video_has_tag.tag_added BETWEEN ? AND
?";
```

Kuva 5.17. *Palvelimen tietokantakysely, säätagien kyselyn toinen osa.*

Tämä kysely (kuva 5.17.) on lähes sama kuin kuvassa 5.14. Ainoana erona on ajan määre. Tässä kyselyssä tagit, jotka hyväksytään tuloksiin, on oltava lisätty aikamääreen sisällä. Aikamääreenä tässä on aika, kun uusi video lähetetään palvelimelle ja siitä kaksi tuntia taaksepäin. Myös tämä kysely on toteutettu silmukan sisällä kuten kuvassa 5.13. on toteutettu geotagien kyselyn toinen osa. Tämä johtuu siis siitä, että edellisen kyselyn kaikki tulokset halutaan tarkastaa. Myös nämä tulokset laitetaan jonoon. Tuloksissa saattaa ilmentyä samoja arvoja, kuin geotagi kyselyssä. Nämä johtuvat siitä, että lähes samalla paikalla saatetaan kuvata lyhyen ajan sisään monta videota. Nämä mahdolliset duplikaatit kuitenkin poistetaan ennen kuin tuloksia syötetään clientille. Tämä duplikaattien poisto tapahtuu myöhemmässä vaiheessa.

Tietokannasta tehtävä seuraava kysely hakee paikkoja, jotka sijaitsevat lähellä kuvauspaikkaa. Tämä kysely on toteutettu samalla tavalla kuin muutkin paikkatietoon liittyvät kyselyt. Seuraavassa kuvassa näkyy kyselyn toteutus.

```
4056      $Kysely = "SELECT name, gp_id FROM geo_places WHERE lat
BETWEEN ? AND ? && lng BETWEEN ? AND ? ORDER BY gp_id DESC LIMIT 6";
```

Kuva 5.18. *Palvelimen tietokantakysely, paikkojen nimien haku.*

Kysely (kuva 5.18.) siis hakee "geo_places" nimisestä taulusta paikan nimen ja id:n. Paikan sijainti on rajattu kuten muissakin paikkatietoon liittyvissä kyselyissä. Tässä kyselyssä paikat, jotka ovat noin 100 metrin päässä kuvaus paikasta tulevat mukaan tagiehdotuksiin. Tulokset on rajattu myös tässä tapauksessa. Tämä toimenpide tehdään, ettei tagikirjasto paisuisi liian suureksi. Tietokannan "geo_places" tauluun lisätään paikkoja xml-tiedostosta. Tämä tiedosto on ladattu <http://geo-position.com> – sivustolta. Sivustolla voi kuka tahansa palveluun kirjautunut käyttäjä lisätä paikkoja sekä niiden koordinaatteja. Sivusto on tarkoitettu pääasiassa navigaattoreille, mutta tietoja voidaan käyttää myös muiden sovellusten kanssa. Ladattavissa oleva tiedosto sisältää muun muassa paikkojen nimiä, yrityksiä ja nähtävyyksiä. Sivustolta voi ladata monen eri maan paikkatietoja. Tässä työssä on käytetty ainoastaan Suomen paikkoja.

```

10 <wpt lat="61.504354000" lon="23.744653000">
11   <name>Särkänniemi</name>
12   <cmt>Särkänniemi Huvipuisto tampereella-Tampere</cmt>
13   <desc>Särkänniemi Huvipuisto tampereella-Tampere</desc>
14 </wpt>
15 <wpt lat="61.497976000" lon="23.750939000">
16   <name>Pääkirjasto Metso</name>
17   <cmt>Pääkirjasto Metso Tampereen pääkirjasto hämeenpuiston
varrella.-Tampere</cmt>
18   <desc>Pääkirjasto Metso Tampereen pääkirjasto hämeenpuiston
varrella.-Tampere</desc>
19 </wpt>
20 <wpt lat="61.493539000" lon="23.763459000">
21   <name>Ratinan stadioni</name>
22   <cmt>Ratinan stadioni Ratinan urheilustadioni-Tampere</cmt>
23   <desc>Ratinan stadioni Ratinan urheilustadioni-Tampere</desc>
24 </wpt>
25 <wpt lat="61.498308000" lon="23.760569000">
26   <name>Keskustori</name>
27   <cmt>Keskustori Tampereen keskus. Tästä kaupunki laajenee jokaiseen
ilmansuuntaan.-Tampere</cmt>

```

Kuva 5.19. Palvelimen tietokantaan ladattavan xml-tiedoston sisältö.

Ladattavan xml-tiedoston rakenne nähdään kuvassa 5.19. Tiedosto sisältää paikan koordinaatit, paikan nimen sekä kuvauksen. Tämä tiedosto päivittyy miltei joka päivä ja sen voi ladata koska tahansa palvelimen tietokantaan. Tietokantaan lataamista varten on kehitetty oma koodi, joka tallentaa tiedot automaattisesti oikeaan paikkaan tietokantaan. Sen toiminta on erittäin helppoa. Tietojen tallentamista varten ei tarvitse, kuin ajaa luotu koodi.

```

2 include("/var/www/functions.php");
3 include("/var/www/database.php");
4
5 global $mysqli;
6 MySQLConnect();

```

Kuva 5.20. Palvelin, tietokantaan tallennetaan xml-tiedosto.

Tiedoston alussa otetaan käyttöön "functions.php" ja "database.php", kuten kuvasta 5.20. on nähtävissä. Nämä ottavat nimensä mukaisesti tietokannan ja funktiot käyttöön, kuten ne on määritelty kyseisissä tiedostoissa. Tämän jälkeen ajetaan komento "global \$mysqli", mikä valmistelee yhteyden tietokantaan. Tämä komento on pakollinen aina, kun tietokantaan tallennetaan tai sitä luetaan. Sen yhteyteen kuuluu myös komento "MySQLConnect()" joka viimeistelee yhteyden tietokantaan.

```

8 set_time_limit(300);

```

Kuva 5.21. Palvelin, tietokantaan tallentamisen esimäärittelyt.

Kuvassa 5.21. asetetaan tietokantakyselylle aika, kuinka kauan se käy ohjelmaa läpi. Tämä on tehtävä, koska tiedosto on todella suuri ja sitä ei ehdi käymään läpi siinä ajassa, joka on kyselyiden oletusarvona. Käsky `set_time_limit(300)` käy käskyä läpi 5 minuuttia. Tämän on katsottu riittävän, jos tietokantaa päivittää riittävän usein, eikä anna tiedoston kasvaa liian suureksi.

```

9  $xmlDoc = new DOMDocument();
10 $xmlDoc->load(
    "http://movie.pori.tut.fi/saxophone/antti/Finland_2009-10-19.xml");

```

Kuva 5.22. *Xml-tiedoston lukemisen aloittaminen.*

Xml-tiedosto (katso kuva 5.19.) luetaan ensin läpi käyttäen komentoa `DOMDocument()` (kuva 5.22.). DOM (Document Object Model) on alusta- ja ohjelmointi kieliriippumaton rajapinta, joka mahdollistaa ohjelmien ja koodien pääsyn sisältöön. Se mahdollistaa myös tietyn dokumentin rakenteen ja tyylin päivittämisen. [82] Tämän jälkeen tiedosto ladataan käyttöä varten.

```

12 $i = 0;
13 foreach ($xmlDoc->getElementsByTagName('wpt') as $node) {
14     $place_name = utf8_decode($node->getElementsByTagName('name')->
        item(0)->nodeValue);
15     $place_desc = utf8_decode($node->getElementsByTagName('desc')->
        item(0)->nodeValue);
16     $place_lat = $node->getAttribute('lat');
17     $place_lon = $node->getAttribute('lon');

```

Kuva 5.23. *Palvelin, xml-tiedoston tietojen luku.*

Kun tiedosto on tallennettu, siitä aletaan etsiä tiettyjä työn kannalta tärkeitä elementtejä. Nämä elementit ovat tässä työssä:

- `name`, tietokantaan tallennettavan paikan nimi
- `desc`, tietokantaan tallennettavan paikan nimen kuvaus
- `lat` ja `lon`, tietokantaan tallennettavan paikan paikkatiedot

Aluksi määrätään minkä tagien sisältä haetaan tietoa. Kuten kuvasta 5.23. on havaittavissa, kaikki halutut arvot sijaitsevat `<wpt>`-tagien sisällä. Tästä johtuen alussa määritellään, että jokaisen `<wpt>`-tagien sisältä luetaan arvot ja ne tallennetaan tilapäisesti muuttujaan `$node`. Tämän jälkeen voidaan tarkentaa mitä muuttujia halutaan tallentaa tietokantaan. Kuvassa 5.23. rivillä 14 ja 15 haetaan jokainen `"name"` ja `"desc"` `<wpt>`-tagien sisältä. Nämä ovat siis paikkojen nimiä ja niiden kuvauksia. Kyseisiin arvoihin täytyy käyttää komentoa `"utf8_decode"`, koska muutoin tiedostosta tulevat kirjaimet eivät näy oikein tietokantaan tallentaessa. `Utf8_decode()`-funktio dekodaa UTF-8 muotoisen sanan noudattamaan ISO-8859-1 formaattia. [83] Kuvassa 5.23. rivillä 16 ja 17 otetaan arvot hieman eritavalla johtuen xml-tiedoston muodosta. Koska `lat` ja `lon` arvot ovat `<wpt>`-tagien sisällä, kuten kuvasta 5.19. on nähtävissä, pitää arvot lukea attribuutin avulla.

```

19      $Kysely = "SELECT gp_id FROM geo_places WHERE lat=? AND lng=? AND
      name=?";

```

Kuva 5.24. Palvelin, xml-tiedoston tarkastelu.

Kuvassa 5.24. näkyvä kysely sijaitsee kuvassa 5.23. alkavan foreach - silmukan sisällä. Kyseinen rakenne johtuu siitä, että tietokantaan tallennetaan ainoastaan yksi arvo kerrallaan ja foreach – silmukka käy xml-tiedoston arvot yksi kerrallaan läpi. Tämä kysely tarkastaa onko kyseinen arvo jo tietokannassa. Jos tieto on tietokannassa, kyseisellä kierroksella ei tehdä mitään toimenpiteitä.

```

27      if (!$riveja > 0) {
28          $mysqli->query("INSERT INTO geo_places (lat, lng, name,
      description) VALUES ('".$place_lat."', '".$place_lon."', '".$
      $place_name."', '".$place_desc."')");
29      } else {

```

Kuva 5.25. Palvelin, xml-tiedoston tietojen syöttö tietokantaan.

Jos tieto ei ole tietokannassa, eli kuvan 5.25. rivin 27 määrittely käy toteen, syötetään kyseiset arvot tietokantaan. Tietokantaan on luotu "geo_places" niminen taulu, jossa on arvoina lat, lng, name ja description. Kyseisellä kierroksella käsiteltävänä olevat arvot syötetään näiden muuttujien tilalle, jos on päästy tähän funktioon. Kuvassa 5.25. näkyvä "else" lauseke tulostaa muussa tapauksessa ainoastaan ennalta määrätyn tiedon käyttäjälle, "Tieto on jo taulussa".

Kun koko funktio on käyty läpi, tulostetaan näytölle "Done!". Tästä tiedetään arvojen tallentuneen tietokantaan oikein ja että koko xml-dokumentti on käyty läpi. Tämä koodi mahdollistaa saman tiedoston läpi käymisen moneen kertaan, koska samaa arvoa ei tallenneta tietokantaan, kuin ainoastaan yhden kerran. Tästä johtuen xml-tiedosto voidaan ladata koska tahansa Internet-sivustolta ja tallentaa se tietokantaan. Tällä hetkellä tietokannan "geo_places" – taulu sisältää hieman alle 30 000 paikan nimeä kuvauksineen ja paikkatietoineen koko Suomen alueelta. Kun tiedosto on käsitelty loppuun, täytyy tietokanta yhteys sulkea. Tämä tapahtuu komennolla "mysqliDisconnect()".

Kaikkien tarpeellisten tietojen tietokannasta haun jälkeen, niille on tehtävä tarkastelu, ettei tagiehdotuksiin tule monta kertaa täysin samaa tagia. Tämä tarkastelu toteutetaan vertaamalla kaikkia tietokantakyselyiden tuloksena tulleita jonoja toisiinsa. Jonoissa verrataan tagin id:tä toisiinsa, koska tagi sanat saattavat olla samoja, tarkoittaen kuitenkin eri asiaa. Tämä johtuu suomen kielen taivutusmuodoista ja niitä ei oteta tässä työssä huomioon.

```

4074     $Osumat = array();
4075
4076     for ($j = 0; $j < count($geo); $j++) {
4077         $Lisaa = true;
4078         for ($k = 0; $k < count($Osumat); $k++) {
4079             if ($Osumat[$k][0] == $geo[$j]['tag_id']) {
4080                 $Lisaa = false;
4081             }
4082         }
4083         if ($Lisaa) {
4084             array_push($Osumat, array($geo[$j]['tag_id'], $geo[$j]
4085                                     ]['tag_name']));
4086         }

```

Kuva 5.26. Palvelin, tietokantakyselyiden tulosten yhdistäminen ja duplikaattien poistaminen.

Aluksi määritellään muuttujan muoto. Koska tässä tallennetaan tulokset jonoon, määritellään muuttuja ”\$Osumat” jonoksi ”**array()**”. Tämän jälkeen tarkastellaan tietokantakyselyn tuloksia jonoista alkio kerrallaan. Kuvassa 5.26. tarkastellaan geotagien tietokantakyselyn tuloksia. Kuten muissakin tapauksissa jonoa lähdetään käsittelemään ensimmäisestä alkioista ja jonon pituus lasketaan, jotta tiedetään mikä on jonon viimeisen alkion järjestysnumero. Rivillä 4077 määritellään muuttuja ”\$Lisää”, jonka avulla tarkastellaan onko syötettävä arvo jo jonossa. Rivillä 4078 tarkastellaan jonon ”\$Osumat” alkioden määrää. Seuraavalla rivillä tarkastellaan sisältääkö jono ”\$Osumat” jo kyseisen arvon, jota sinne yritetään syöttää. Jos jonon sisältämän muuttujan ”\$Lisää” arvo on epätosi, kyseistä arvoa ei lisätä jonoon. Jos taas arvoa ei ole vielä jonossa, eli rivin 4083 ehtolause täyttyy, lisätään jonoon ”\$Osumat”, tässä tapauksessa, jonon ”\$geo” tagin id sekä nimi.

Tämä tarkastelu käydään läpi siis kaikille tuloksille ja lopuksi saadaan vastaukseksi jono ”\$Osumat”, joka sisältää uniikit tagien nimet sekä niiden id:t. Tagin id on otettava aina mukaan, koska muutoin sama tagi tallennettaisiin aina uudestaan tietokantaan. Tämä pyritään estämään juuri id:n avulla.

Tästä toteutuksesta eroaa ”geo_places” -taulu. Kun kyseisestä taulusta lisätään paikkojen nimiä niiden id:t eivät ole sidoksissa muihin tagien id:n millään tavalla. Tämän vuoksi tagijonoon lisääminen on jouduttu toteuttamaan eri tavalla. Kun taulusta saadut vastaukset lisätään tagiehdotuksiin, asetetaan niiden id:t nolllaksi. Jos puhelimesta valitaan kyseinen tagi kuvastamaan videota, niin tieto lähetetään palvelimelle uutena tagina. Tällöin palvelin lisää tagin omaan kirjastoonsa ja asettaa automaattisesti id:n, joka on tietokannan tauluihin sopiva.

```

4121     $xml = "<tags>";
4122     if ($tag != null) {
4123         for($i=0; $i<count($Osumat); $i++){
4124             $xml .= "<tag id='". $Osumat[$i][0]. "'>". $Osumat[$i][1]
4125         ]. "</tag>";
4126     }
4127     } else {
4128         $xml .= "<tag id='10'>mobile</tag>";
4129     }

```

Kuva 5.27. Palvelin, tietojen syöttö *demo.php*-tiedostoon.

Kun tulokset on käsitelty ja duplikaatit poistettu, tehdään tuloksista xml-tiedosto. Tiedoston luonti aloitetaan kuvassa 5.27. rivillä 4121. Tässä luodaan muuttuja ”\$xml”, jonka sisään kaikki arvot syötetään. Tiedosto aloitetaan tagilla ”<tags>”. Tämän jälkeen tehdään tarkastelu onko muuttuja ”\$tag” nolasta poikkeava. Eli jos tässä muuttujassa on arvoja, niin voidaan tietojen syöttö omaan xml-tiedostoon aloittaa. Rivillä 4123 tarkastetaan, montako arvoa tulee syöttää tuloksiin. Tarkasteltava jono on ”\$Osumat”, johon kaikki tagiehdotukset on syötetty, kuten kuvasta 5.27. on nähtävissä. Kyseinen jono käydään siis läpi alkoi kerrallaan ja alkiot syötetään ”<tag>” – tagien sisään. Rivillä 4124 merkintä ”\$xml .=” tarkoittaa saman xml-tiedoston jatkamista. Tähän tiedostoon lisätään tagien id:t ja tagin nimet. Ennen kuin tuloksia laitetaan xml-tiedostoon, tulokset ovat ainoastaan jonossa. Tällainen jono näyttää seuraavanlaiselta (kuva 5.28.):

```

1  Array
2  (
3      [0] => Array
4          (
5              [0] => 524
6              [1] => Aika
7          )
8
9      [1] => Array
10         (
11             [0] => 523
12             [1] => Toimiiko
13         )
14
15     [2] => Array
16         (
17             [0] => 520
18             [1] => Esittelijät
19         )
20

```

Kuva 5.28. Tagien tulostus esimerkki selaimessa, tulokset jonossa. Tästä luodaan xml-tiedosto.

Jos näitä tietoja ei jostain syystä ole saatavilla, siirrytään suoraan else – lausekkeeseen ja annetaan tagiehdotukseksi ainoastaan sana ”mobile” jonka id on

määritelty käsin. Tällainen arvo on tietokannassa, eli tagin id:tä ei ole keksitty päästä. Tähän lausekkeeseen voidaan joutua, jos tietokannassa on jokin virhe tai, jos käyttäjä on uusi ja kuvaa ensimmäistä videotaan eikä ole syöttänyt vielä tageja tietokantaan eikä puhelimen gps toiminto ole käytössä. Toisin sanoen on melko vähän tilanteita jolloin palvelin ei ehdota tagiehdotuksiksi mitään.

```
4130     $xml .= "</tags><title>". $vd_name. "</title><desc>".  
      $vd_description. "</desc>";  
4131     return $xml;
```

Kuva 5.29. Palvelimen xml-tiedoston luominen, kuvaus- ja otsikkotiedot.

Tässä on nähtävillä palvelimella luotavan xml-tiedoston loppu osa (kuva 5.29.). Xml-tiedostoon siis syötetään vielä otsikko sekä kuvaus, jotka on haettu tietokannasta aikaisemmin (kuva 5.9.). Kun kaikki tiedot on syötetty tietokantaan, annetaan komento ”**return** \$xml”, joka palauttaa ja kokoaa juuri luodun xml-tiedoston. Tämä mahdollistaa xml-tiedoston muiden funktioiden käyttöön.

```
116     $xml = demoSuggestTags($usr_id, $vd_id, $lat, $lng);  
117     echo "<res><vid>". $vd_id. "</vid>". $xml. "</res>";
```

Kuva 5.30. Palvelimen xml-tiedoston palauttaminen client -ohjelmiston käyttöön.

Tässä palautetaan luotu xml-tiedosto client -ohjelmiston käyttöön. Tulokset toisesta tiedostosta sijoitetaan tässä lauseessa kohtaan ”</vid>”. \$xml. ”</res>”. Jonka jälkeen xml-tiedosto on valmiina client -ohjelmiston toimia varten. Xml-tiedosto näyttää tässä tulostuksessa seuraavanlaiselta.

```
34     <res><vid>1526</vid><tags><tag id='524'>Aika</tag><tag id='523'>  
      Toimiiko</tag><tag id='520'>Esittelijät</tag><tag id='521'>ständi  
      </tag><tag id='522'>meininkiä</tag></tags><title>Loikka  
      </title><desc>Tagi testi </desc></res>
```

Kuva 5.31. Tagien tulostus esimerkki selaimessa, tuloksista luotu xml-tiedosto.

Xml-tiedosto hakee ensin videon id:n, joka on käsittelyssä parasta aikaa. Tämän jälkeen tulevat tagiehdotukset. Tagiehdotuksien mukana on tagin uniikiksi tekevä id. Tagien jälkeen tulee ehdotus otsikko, mikä on sama kuin käyttäjän viimeksi syöttämä otsikko. Lopuksi tulee kuvaus videolle, joka on ehdotuksena sama, kuin viimeksi syötetty kuvaus.

Client -ohjelmisto käy lukemassa tämän tiedoston ja saa sieltä tagiehdotukset, jotka palvelin on hakenut tietokannasta. Client -ohjelmisto lukee tiedoston puhelimeen tietystä tiedostosta. Tiedosto, jota client -ohjelmisto käsittelee, on samanlainen kuin kuvassa 5.31.

Client käy lukemassa tiedoston ”demo.php”, joka sisältää juuri luodun xml-tiedoston. Aluksi luodaan yhteys palvelimeen, jonka jälkeen määrätään, mitä tiedostoa kyseisestä paikasta käsitellään.

Seuraavaksi tarkastellaan onko videotiedoston lataus palvelimelle onnistunut. Jos videotiedoston lataus on onnistunut, tulee käyttäjälle matkapuhelimeen ilmoitus ”Uploading OK”. Tämän jälkeen client –ohjelmisto tarkastelee demo.php –tiedostoon ladattuja tietoja. Client määrittää missä kohdin tiedostossa haluttu tieto sijaitsee ja lukee kaiken tarvittavan muistiin myöhempää käsittelyä varten. Client käy lukemassa demo.php –tiedostosta videon id:n, otsikko- ja kuvausehdotukset sekä tagiehdotukset. Jos tagien välistä ei löydy tietoja on lataaminen epäonnistunut. Jotta tällaista ei pääse tapahtumaan, palvelin lähettää aina vähintään ”mobile” – tagin käyttäjälle, kuten jo aikaisemmin selvitettiin (kuva 5.27.).

Client –ohjelmistoon on myös rakennettu merkkitarkastelu. Tämä on välttämätön suomenkielessä, koska Python, jolla client on rakennettu, ymmärtää ainoastaan englantilaisen merkistön.

Kun client on hakenut kaikki halutut tiedot, näytetään ne matkapuhelimen näytöllä. Käyttäjän lisätessä kontekstitietoa videosta, tulostetaan otsikkoa ja kuvausta lisätessä palvelimelta saadut ehdotukset tekstikenttiin. Tagiehdotukset tulostetaan puhelimeen listana, koska niitä on yleensä hieman enemmän. Tagiehdotuksia tulee automaattisesta kontekstitiedosta ja käyttäjän tagi historiasta riippuen 6-20kpl. Tämän on havaittu olevan melko kattava määrä, koska tageja ollessa enemmän käyttäjä kyllästyy listan selaamiseen. Jos tageja on taas vähemmän, käyttäjä ei välttämättä löydä itselleen sopivia tageja listasta.

5.4.2. Tulevaa MoViE -palveluun

Palvelimelle on luotu edellä kuvattujen toimintojen lisäksi vielä käyttöönottamattomia ominaisuuksia. Toteutukseen vaadittavat koodit on jo rakennettu valmiiksi, mutta niitä varten ei ole vielä luotu kaikkia osia valmiiksi. Tässä kappaleessa esitellään, mitä tagiehdotuksia lisääviä toteutuksia on tulossa MoViE -palveluun.

Yksi ehdotuksista on paikkakunnan tapahtumat. Tämä koodi pyytää ensin paikannimen, joka tulee automaattisyytteenä. Kun paikan nimi on saatu, koodi lukee ”www.tapahtuma.tv” –sivuston lähdekoodin. Lähdekoodista pyritään erottelemaan kaikki maksulliset tapahtumat, kuten näytökset ja urheilutapahtumat, ja nämä tulevat ehdotuksiksi puhelimeen. Tällainen on ajateltu olevan hyvä, koska erilaisissa tapahtumissa kuvataan yleensä paljon videoita ja MoViE -palvelu on suunnattu muun muassa tällaiseen käyttöön (vertaa case I ja case II).

Toinen palvelimelle integroimaton tagiehdotuksiin perustuva sovellus on muiden käyttäjien tagien vertailu. Tämä vertaa palvelimelle lähetettyä tagia ja etsii sitä kaikista muista videoista. Jos tagi löytyy jostain muusta videosta, lähetetään kyseisen videon kaikki tagit käyttäjälle. Tätä käytetään lähinnä sen vuoksi, että eri käyttäjät kuvaavat samantyyllisiä videoita. Näin tagiehdotuksiin voisi tulla hyviä ehdotuksia ja käyttäjän ei tarvitsisi miettiä tätä niin paljoa. Tämän ongelmana on, että käyttäjät voivat käyttää täysin samaa tagia tarkoittaen eri asiaa.

Näiden lisäksi on rakennettu lähinnä matkapuhelinta ajatellen ohjelma, joka näyttää kartalta 20 viimeisimmän videon kuvauspaikan. Ohjelma kysyy käyttäjältä, mitä kaupunkia tämä haluaa tarkkailla ja miten tarkkaan (kuinka läheltä). Kartta on rakennettu kuvista, jotta se toimisi mahdollisimman monessa matkapuhelimessa. Palvelu on suunniteltu liitettävän MoViE –sivustoon.

Tagiehdotuksiin kuulumaton palvelu on myös suunnitteilla MoViE –palveluun. MoViE:n pyritään saamaan tuki reaaliaikaisen videon lähetykseen ja katseluun. Tämä tapahtuisi verkon yli live streamina.

5.5. Sovellusten yhteensopivuus

MoViE –palvelu on luotu siten, että sen upottaminen on mahdollista myös muihin sivustoihin. Upottaminen vaati joissain tapauksissa pieniä muutoksia, mutta se onnistuu suhteellisen pienellä työllä. Upottamisen voi toteuttaa siirtämällä koko sovelluksen, eli tietokannat ja kaikki siihen liittyvät tiedostot, toimimaan toiselle palvelimelle. Toinen vaihtoehto on linkittää muualla toimiva sivusto jo käytössä olevaan MoViE –palvelimeen. MoViE –palvelu olisikin varsin oiva sovellus esimerkiksi lehden toimittajille, koska he voisivat kuvata jutun paikanpäällä ja lähettää videon välittömästi palvelimelle, käyttäen MoViE –palvelua niin sanottuna videoarkistona.

MoViE –palvelua on käytetty muun muassa Pori Jazzien kilpailusivustolla. Kilpailusivusto, saxoPHONE, on toteutettu MoViE –palvelun päälle. Sivuston muokkaaminen ei vie aikaa, mutta se vaatii toimiakseen tietokannan sekä PHP tuen.

6. NYKYTILANNE JA TULEVAISUUS

Tulevaisuudessa tagien luontia tultaneen helpottamaan. Tagien käytön ongelma nyt on muun muassa niiden hankala lisääminen. Esimerkiksi puhelimelle tagien lisääminen on aikaa vievää, joten ihmiset eivät monesti jaksakaan lisätä niitä. Eräs ratkaisu tähän on saneltavat tagit. Puhelimesta ei vielä välttämättä riitä teho erotella puheesta sanoja, joten tämä tulisi toteuttaa palvelimella.

Cherubini et al. ovat vertailleet kuviin tekstillä ja puheella liitettäviä tageja. Heidän tutkimuksessaan puheella lisättävät tagit ainoastaan tallennettiin puhelimen muistiin, eikä niitä yritetty muuttaa tekstiksi, kuten tulevaisuudessa todennäköisesti tehdään. Tutkimuksessa vertailtiin kuinka hyvin tagit muistettiin ja miten paljon tageja lisättiin. Tutkimuksessa todettiin, ettei puheella lisättävissä tageissa ole vielä huomattavaa hyötyä tekstillä lisättäviin tageihin. Puheella lisättäviä tageja ei haluttu käyttää, jos oltiin julkisella paikalla. Tekstillä lisättävät tagit olivat taas hankalia, kun oltiin liikkeessä. Kuitenkin tekstillä lisätyt tagit oli helpompi muistaa kuvista. [84]

Tekniikan kehittyessä myös kuvien tunnistus taggaamisen apuvälineenä saattaa yleistyä. Koska videot koostuvat todella monesta kuvasta, tällainen toiminta vaatii paljon tehoa. Tällainen toiminta tullee toteuttamaan tehokkaalla palvelimella, eli toiminta vaatii Internet-yhteyttä. Toiminta vaatii myös paljon älyä, jottei se jatkuvasti yritä tunnistaa kuvia, joiden kohteena on sama esine tai asia. Tällainen ominaisuus on jo toteutettu kuville. Toteutuksessa on tallennettu tietokantaan paljon erilaisia kuvia. Ohjelma vertaa otettua kuvaa tietokannassa oleviin kuviin ja yrittää etsiä näistä riittävästi yhtäläisyyksiä tunnistukseen kuvan. Kuvan tunnistus voidaan myös toteuttaa menetelmillä ja algoritmeilla, joiden avulla datasta pyritään tunnistamaan säännönmukaisuuksia. Toistaiseksi oleellinen este palvelun toteuttamisessa esimerkiksi MoViE –palveluun on tunnistamisen monimutkaisuus. Sovellus voitaisiin toteuttaa siten, että vertaillaan käyttäjän ottamia videoita tai kuvia juuri otettuun kuvaan tai videoon. Mikäli näistä löytyy riittävästi yhtäläisyyksiä, annettaisiin automaattisesti edellisellä kerralla syötetyt tagit uudestaan ehdotuksina käyttäjälle. [85; 86]

Tulevaisuudessa myös tagikirjastolta vaaditaan huomattavasti enemmän. Se saattaa olla palvelimella Internetissä, josta sitä voidaan käyttää. Tagikirjasto on yhteensopiva kaikkien sovellusten kautta. Lisäksi voisi olla erilaisiin paikkoihin upotettavia tageja. Esimerkiksi jos otetaan kuva juomatölkistä, siihen olisi valmiiksi upotettu valmistajan nimi, joka tulee tagiehdotuksena käyttäjälle ja jää myös henkilökohtaiseen tagikirjastoon. Tagikirjastosta voisi myös kehittää ”opetetun”

tagikirjaston. Tässä esimerkiksi matkapuhelin voisi ottaa erilaista kontekstietoa ympäristöstä ja ehdottaa sen mukaan eri tageja sekä samalla rajata jotkin tagit täysin pois sopimattomuutensa vuoksi. Lisäksi tageja voisi myös hakea erilaisten palvelujen kautta reaaliaikaisesti, eli rakentaa mittava tekoäly tagiehdotusten taakse, kuten Harri Ketamo artikkelissaan esittelee. [30]

7. YHTEENVETO

Työssä toteutettiin tagikirjasto MoViE –palveluun. Tagikirjaston tarkoituksena oli palvella matkapuhelimessa toimivaa client –ohjelmistoa avainsanojen lisäyksessä videoon. Tagiehdotusten antamisessa käytetään matkapuhelimelta saatua kontekstietoa. Tagikirjastosta tuli melko kattava käyttäjälle, jolla on mahdollisuus saada paikkatiedot. Tagikirjasto riittänee avustavana tagiehdotusten antajana useimmille käyttäjille. Kirjasto parantuu huomattavasti, kun tagikirjaston lisäosat integroidaan palveluun.

Kattavan tagikirjaston tekeminen palveluun on kuitenkin hankalaa, jos käyttäjiä ei ole riittävästi lisäämään tagisanoja kirjastoon. Jos tagikirjasto olisi laajemmassa käytössä, sen tagiehdotukset parantuisivat. Ohjelmisto on tällä hetkellä prototyyppi, mikä rajoittaa käyttäjien määrän muutamaa kymmeneen. Tämä ei lisää tagikirjaston ehdotuksia halutulla tavalla tai laadulla.

Tagikirjaston tagiehdotukset ovat kattavia, jos saadaan paikkatiedot. Ilman paikkatietoja tagikirjasto lähettää käyttäjälle ainoastaan käyttäjän viisi viimeisintä tagia. Käyttäjän ollessa esimerkiksi kaupungissa saadaan huomattavasti parempia tagiehdotuksia, kuin harvaan asutuilla alueilla. Tämä johtuu kaupungissa olevien tagattavien asioiden määrästä. Lisäksi käyttäjiä on enemmän lisäämässä tagisanoja melko samassa paikassa.

Tagikirjaston kehittäminen vaatii lisätutkimusta. Vaikka käyttäjiä ei olisi vielä paljoa, voidaan tageja generoida erilaisilla hauilla Internetistä. Hakuja voidaan tehdä joko matkapuhelimesta saatavilla kontekstiedoilla tai syötetyillä tagisanoilla. Jos käytetään käyttäjän syöttämiä tagisanoja, tulisi palvelun olla jatkuvasti tietoinen syötetyistä tagisanoista, jotta niiden avulla tehtävät haut voitaisiin suorittaa välittömästi. Tämä edellyttäisi jatkuvaa Internet-yhteyttä ja älyä sovellukselta. Tässä tapauksessa kyseinen toiminta ei ole kannattavaa. Internetistä haettavat tagisanat voitaisiin hakea eri hakupalvelimilta tai RSS -syötteistä. RSS -syötteistä hakeminen voisi kohdistua esimerkiksi uutisiin, jolloin saataisiin mahdollisimman ajankohtaisia sanoja tagiehdotuksiksi. Tällainen toiminta vaatii vielä paljon kehitystä.

Tagien syöttöä voitaisiin myös nopeuttaa käyttämällä ennustavaa tekstinsyöttöä (T9). Tämä voisi helpottaa ja nopeuttaa matkapuhelimilla lisättäviä tageja. Samalla tekstinsyöttöjärjestelmä kehittäisi matkapuhelimeen toista tagikirjastoa, joka ei olisi riippuvainen Internet-yhteydestä. Tämän avulla ei kuitenkaan saataisi varsinaisia tagiehdotuksia, mutta se auttaisi ongelmassa matkapuhelimilla.

Jos esimerkiksi MoViE-palvelussa olisi suuria käyttäjämääriä, tulisi sinne toteuttaa valvontaa. Valvonnan tarkoituksena olisi poistaa sopimattomat tagisanat

yleisestä kaikkien käytössä olevasta tagikirjastosta. Palvelu tulisi toteuttaa automaattiseksi, koska käyttäjämäärän ollessa suuri myös tagien määrä olisi suuri. Tällainen toiminta tulisi toteuttaa esimerkiksi sanojen vertailulla. Tähän ongelmaan ei olla diplomityössä paneuduttu. On myös huomattava, että tagikirjasto on oltava hyvä siinä mielessä, että väärä taggaus voi pahimmassa tapauksessa johtaa tiedon hukkumiseen suurten tietomäärien joukkoon.

Tagit eivät ole vielä kaikille selviä, eivätkä ne ole käyttäjien suuressa käytössä. Tulevaisuudessa, kun tietomäärät kasvavat yhä suuremmiksi, tietojen löytäminen hankaloituu. Tagien käyttö saattaisi olla ratkaisu ongelmaan. Käyttäjien on kuitenkin vaikea alkaa käyttämään tageja ennen niiden ymmärtämistä. Tageja on melko helppo käyttää verkkoon tallennetuissa tiedoissa, mutta kovalevylle tallennetut tiedot ovat vielä helpon taggauksen ulkopuolella. Kovalevylle tallennettavat tiedot vaativat erillisen ohjelman taggausta varten. Tiedot on tallennettava ainoastaan ”järkevä” tiedostorakenteen avulla.

Lähteet

- [1] Laitinen Janne. GPS-paikkatiedon liittäminen matkapuhelinpalveluihin. DI-työ. Lappeenranta 2004. Lappeenrannan teknillinen yliopisto, Tietotekniikan osasto, Tietoliikennetekniikan laitos. 79p.
- [2] Multisilta Jari, Suominen Marko. MoViE: Mobile Video Experience. MindTrek, Tampere 2009, Lokakuu. Tampere, MindTrek. Pp 157-161.
- [3] Multisilta Jari, Mäenpää Marjo. Mobile Video Stories. Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts, SESSION: Interactive stories 2008. Pp 401-406.
- [4] Nielsen Jakob. WWW suunnittelu (alkup. Designing Web Usability, suom. Timo Haanpää). Jyväskylä 2000, IT Press. 417p.
- [5] Nielsen Jakob, Loranger Hoa. Prioritizing Web Usability. Berkeley 2006, New Riders. 406p.
- [6] Tieto [WWW]. [viitattu 27.07.2009] saatavissa: <http://fi.wikipedia.org/wiki/Tieto>
- [7] Ritva Koskennurmi-Sivonen. Tieto, taito ja tekijän hiljainen tieto [WWW]. [viitattu 27.07.2009] saatavissa: <http://www.helsinki.fi/~rkosken/tietojataito.html>
- [8] Pentti Routio. Tiedon lajeja [WWW]. [viitattu 27.07.2009] saatavissa: <http://www2.uiah.fi/projekti/metodi/048.htm>
- [9] Niiniluoto Ilkka. Informaatio, tieto ja yhteiskunta [PDF]. [viitattu 30.07.2009] saatavissa: www.uta.fi/jour/opiskelu/P1_lukemisto/Niiniluoto2.pdf
- [10] Mallivastaukset 2005 / sosiaalityö [PDF]. [viitattu 31.07.2009] saatavissa: http://www.valt.helsinki.fi/tiedekunta/valinnat/kysymykset05/arv_sostyo.pdf
- [11] Otto Lappi. Descartes [WWW]. [viitattu 31.07.2009] saatavissa: http://www.mv.helsinki.fi/home/olappi/lukionfilosofia/filosofit/descartes/descartessin_epailyn_metodi.htm
- [12] Hiram Caton. Kennington on Descartes' Evil Genius [WWW]. [viitattu 03.08.2009] saatavissa: <http://www.jstor.org/stable/2708894?seq=1>
- [13] Tiedon käsite [WWW]. [viitattu 03.08.2009] saatavissa: <http://www.uta.fi/viesverk/johdviest/tietoinformaatio/tietokasite.html>
- [14] Järviniitty Juho. Kontekstitietoiset liikkuvat web-palvelut. Tutkielma. Helsinki 2005. Helsingin yliopisto, Matemaattis- luonnontieteellinen osasto. 19p.

- [15] IDC Predicts Increased Number of Digital Pictures Snapped by Affluent PC Users Will Boost Global Photo Printing [WWW]. [viitattu 08.12.2009] saatavissa:
http://www.eetimes.com/press_releases/bizwire/showPressRelease.jhtml?articleId=413980&CompanyId=2
- [16] Lehtinen Juha, Aaltonen Antti, Huuskonen Pertti, Salminen Ilkka. Personal Content Experience. West Sussex England 2007, Wiley.
- [17] Dey Anind K. Understanding and Using Context. Personal and Ubiquitous Computing 5 (2001) 1, pp. 4-7.
- [18] Schilit Bill N., Adams Norman, Want Roy. Context-Aware Computing Applications. Workshop on Mobile Computing Systems and Applications, Santa Cruz, December, 1994. Washington, IEEE Computer Society. Pp 85-90.
- [19] Albrecht Schmidt, Michael Beigl and Hans-W. Gellersen. There is more to Context than Location. Computers & Graphics Journal, Elsevier, 23 (1999) 6. Pp 893-902.
- [20] Pascoe Jason. Adding Generic Contextual Capabilities to Wearable Computers. Proceedings of the 2nd IEEE International Symposium on Wearable Computers, 1998 pp. 92-99.
- [21] Niinivaara Olli. Idean kontekstia kuvaava metatieto. Pro gradu. Helsinki 2002. Helsingin yliopisto, Tietojenkäsittelytieteen laitos. 82p.
- [22] Baca Murtha. Introduction to Metadata: Pathways to digital information. Los Angeles 2000, Getty Publications. 96p.
- [23] Lehtinen, A., Salminen, A., Nurmeksela, R. Metatiedot organisaatioiden sisällönhallinnassa. Metatiedot suomalaisen lainsäädäntöprosessin tiedonhallinnassa, Eduskunnan kanslian julkaisu (2005) 7, 4-13.
- [24] Gilliland-Swetland A.J. Setting the state: Defining metadata 2000 (kirjassa: Baca M. Introduction to Metadata: Pathways to digital information).
- [25] Perttu Tolvanen. Www-sisällönhallinta ja www-sisällönhallintajärjestelmien ominaisuudet. Kandidaatintutkielma. Jyväskylä 2006. Jyväskylän yliopisto, Tietojenkäsittelytieteiden laitos. 38p.
- [26] Jeskanen Kirsi, Naumanen Paula, Suominen Nina, Suominen Saila [WWW]. [viitattu 10.12.2009] saatavissa:
<http://www.dipoli.tkk.fi/tietokoulutus/INFO/tiivistelmat/20057.html>
- [27] Lauri Aalto. Bluetooth-paikannus Symbianilla. Sivuainetutkielma. Oulu 2004. Oulun yliopisto, Tietojenkäsittelytieteiden laitos. 39p.
- [28] Aittola Markus. Paikkatietoiset palvelut pienpäälaitteissa. Diplomityö. Oulu 2003. Oulun yliopisto, Sähkö- ja tietotekniikan osasto. 72p.

- [29] Nurmi Petteri. Identiteettiteoriaan ja kontekstietoon perustuva käyttäjän-mallinnus. TIKI aine. Helsinki 2004. Helsingin yliopisto, Tietojenkäsittelytieteen laitos. 13p.
- [30] Ketamo Harri. Suffering from information overload? Try Teachable Media Agents. Online Educa Berlin 2009, 2-4.12.2009, Berlin, Germany.
- [31] Smith Gene. Tagging: People-Powered Metadata for the Social Web. Berkeley 2007, New Riders. 216p.
- [32] Bäck Asta, Melin Magnus, Näkki Pirjo, Vainikainen Sari, Sarvas Risto, Seppälä Lassi, Vihavainen Sami. Tags and tagging - Creating meanings, organising, and socialising with metadata. Espoo 2008, VTT, 2449. 86p.
- [33] Näkki Pirjo. Käyttäjäkokemuksen suunnittelu semanttiseen mediapalveluun – Tarkastelussa kouluretkien tarinat. Diplomityö. Helsinki 2006. Teknillinen korkeakoulu, Tietotekniikan osasto. 101p.
- [34] Kajas Kimmo. Ontologia: katteeton lupaus vai todellinen työkalu?. Pro gradu. Tampere 2008. Tampereen yliopisto, Tietojenkäsittelytieteiden laitos. 68p.
- [35] Laaksonen Salla-Maaria. Organisaatiomaineen tarinat sosiaalisen median keskusteluissa. Pro gradu. Helsinki 2008. Helsingin yliopisto, Valtiotieteellinen tiedekunta. 86p.
- [36] Marlow Cameron, Naaman Mor, Boyd Danah, Davis Marc. HT06, tagging paper, taxonomy, Flickr, academic article, to read. Proceedings of the seventeenth conference on Hypertext and hypermedia, SESSION: Social networks, networking & virtual communities 2006. Pp 31-40.
- [37] Yhteisöllisyys verkossa [WWW]. [viitattu 09.12.2009] saatavissa: <http://mediakompassi.yle.fi/aikuiset/mediailmiot/yhteisot>
- [38] Ontologia (tietojenkäsittelytiede) [WWW]. [viitattu 09.12.2009] saatavissa: [http://fi.wikipedia.org/wiki/Ontologia_\(tietojenkäsittelytiede\)](http://fi.wikipedia.org/wiki/Ontologia_(tietojenkäsittelytiede))
- [39] Folksonomia [WWW]. [viitattu 09.12.2009] saatavissa: <http://fi.wikipedia.org/wiki/Folksonomia>
- [40] Tagien syvemmästä olemuksesta [WWW]. [viitattu 07.08.2009] saatavissa: http://matkalla.org/blog/archives/2008/11/tagien_syvemmasta_olemuksesta.html
- [41] Nokia 5800 XpressMusic puhelimen tuoteseloste [WWW]. [viitattu 26.11.2009] saatavissa: <http://www.nokia.fi/tuotteet/kaikki-puhelimet/nokia-5800-xpressmusic/tarkka-tuoteseloste>
- [42] Poikselkä Piritta. Mobiili sisällön luominen ja julkaisu osana sosiaalisen median palveluja. Diplomityö. Pori 2009. Tampereen teknillinen yliopisto, Tieto- ja sähkötekniikan tiedekunta. 86p.

- [43] Torniai, C., Battle, S. & Cayzer, S. Sharing, Discovering and Browsing Geotagged Pictures on the Web. Digital Media System Laboratory, HP Laboratories Bristol 2007. HPL-2007-73. 19p.
- [44] Hurn Jeff. GPS maailmanlaajuinen satelliittinavigointijärjestelmä (alkup. Trimble navigation, suom. Henttu P., Lehtoranta V.K.). Helsinki 1993, Noviprop Oy. 96p.
- [45] GPS - Global Positioning System [WWW]. [viitattu: 15.01.2010] saatavissa: <http://www.netlab.tkk.fi/opetus/s38118/s98/htyo/8/yleiskuva.shtml>
- [46] Ritakallio Samuel. Suomen pysyvän GPS-verkon uudistaminen GNSS-yhteensopivaksi. Insinööritoimisto. Helsinki 2009. Metropolia Ammattikorkeakoulu, maanmittaustekniikka. 81p.
- [47] Koskinen Kimmo. Paikka- ja tilannetiedon hyödyntäminen sovelluksissa. Diplomityö. Lappeenranta 2005. Lappeenranta teknillinen yliopisto, Tietotekniikan osasto. 56p.
- [48] Jani Hyttinen, Ismo Trast, Markus Huhtinen. GPS ja muut paikannusjärjestelmät [PDF]. [viitattu 10.09.2009] saatavissa: <http://www.ncp.fi/koulutusohjelmat/metsa/PaikkatietoWWW/paikannus/paikannus.pdf>
- [49] Kuva [WWW]. [viitattu 16.01.2010] saatavissa: http://www.luisprada.com/Protected/IMAGES/earth_long-lat.jpg
- [50] Mitä on Geokoodaus [WWW]. [viitattu 10.09.2009] saatavissa: <http://www.karttago.fi/paikkatieto/geokoodaus>
- [51] Geokoodaus verkkojen avulla [WWW]. [viitattu 10.09.2009] saatavissa: <http://www.ncp.fi/koulutusohjelmat/metsa/PaikkatietoWWW/Analyysi/Geokoodaus.htm>
- [52] Valkama Juha. GPS-järjestelmän rakenne ja toimintaperiaate. Insinööritoimisto. Tampere 2006, Tampereen ammattikorkeakoulu, Tietojenkäsittely. 34p.
- [53] Hakolahti Teemu. Mobiilit paikkatietosovellukset. Pro gradu. Joensuu 2003. Joensuun yliopisto, Tietojenkäsittelytiede. 82p.
- [54] Rainio Antti. Henkilökohtainen navigointi - Markkinat, teknologia ja sovellukset. Espoo 2000, VTT, 2037. 124p.
- [55] Standardi lupaa a-gps -ratkaisujen yhteensopivuutta [WWW]. [viitattu 18.12.2009] saatavissa: <http://m.digitoday.fi/?page=showSingleNews&newsID=200511488>
- [56] Kemppi Paul. Database Correlation Method for Multi-System Location. Diplomityö. Helsinki 2005. Teknillinen korkeakoulu, Sähkö- ja tietoliikennetekniikan osasto. 78p.

- [57] Facebook Factsheet [WWW]. [viitattu 11.09.2009] saatavissa:
<http://www.facebook.com/press/info.php?factsheet>
- [58] Microsoft osti palan Facebookia [WWW]. [viitattu 11.09.2009] saatavissa:
<http://www.hs.fi/talous/artikkeli/Microsoft+osti+palan+Facebookia/1135231288136>
- [59] <http://www.facebook.com> [WWW]. [viitattu 12.09.2009] saatavissa:
<http://fi.shvoong.com/internet-and-technologies/portals/1943128-http-www-facebook-com/>
- [60] Googlen Facebook-kehittelyt etenevät [WWW]. [viitattu 12.09.2009] saatavissa:
<http://www.taloussanomat.fi/teknologia/2007/10/31/googlen-facebook-kehittelyt-etenevat/200727298/133>
- [61] Lyyra yrittää tehdä kotimaiset facebookit [WWW]. [viitattu 13.09.2009] saatavissa: <http://www.taloussanomat.fi/markkinointi/2007/10/02/lyyra-yrittaa-tehda-kotimaiset-facebookit/200724302/135>
- [62] Ellison Nicole B., Steinfield Charles, Lampe Cliff. Journal of Computer-Mediated Communication, 12 (2007) 4. Pp. 1143-1168.
- [63] Mäenpää Jenni, Männistö Anssi. Kun kaikki videoivat kaikkea. Juvenes Print 2009, Tampereen yliopisto, Tiedotusopin laitos. 118p.
- [64] Samsung [WWW]. [viitattu 03.12.2009] saatavissa:
http://www.samsung.com/fi/consumer/tv-home-theatre/tv/led-tv/UE40B8050XWXXE/index.idx?pagetype=prd_detail
- [65] Ei ihme, että Youtube ärsyttää kilpailijoita [WWW]. [viitattu 13.10.2009] saatavissa: <http://www.digitoday.fi/viihde/2007/06/29/ei-ihme-etta-youtube-arsyttaa-kilpailijoita/200716082/66>
- [66] YouTube [WWW]. [viitattu 13.10.2009] saatavissa:
<http://www.youtube.com/t/about>
- [67] Google ostaa YouTuben [WWW]. [viitattu 13.10.2009] saatavissa:
<http://www.hs.fi/talous/artikkeli/Google+ostaa+YouTuben/1135222189684>
- [68] Google perusti vararahaston YouTuben mahdollisia syytteitä varten [WWW]. [viitattu 14.10.2009] saatavissa:
<http://www.digitoday.fi/bisnes/2006/11/15/google-perusti-vararahaston-youtuben-mahdollisia-syytteita-varten/200620561/66>
- [69] Suomenkieliset mainokset tulevat Youtube-videoihin [WWW]. [viitattu 14.10.2009] saatavissa:
http://www.tietokone.fi/uutiset/2008/suomenkieliset_mainokset_tulevat_youtub_e_videoihin

- [70] Viddler [WWW]. [viitattu 15.10.2009] saatavissa:
<http://www.viddler.com/learn-more>
- [71] Flickr [WWW]. [viitattu 16.01.2009] saatavissa:
<http://fi.wikipedia.org/wiki/Flickr>
- [72] Flickr [WWW]. [viitattu 01.12.2009] saatavissa: <http://www.flickr.com/about>,
<http://www.flickr.com/help/photos/#18>
- [73] Facebook Lite haastaa Twitterin? [WWW]. [viitattu 02.12.2009] saatavissa:
<http://www.m-brain.com/fi/blog/a/2009/08/12/facebook-lite-haastaa-twitterin>
- [74] The number of people using Twitter in Finland? [WWW]. [viitattu 04.12.2009] saatavissa: <http://tuomaslonka.com/post/266298624/the-number-of-people-using-twitter-in-finland>
- [75] Ahern Shane, Davis Marc, Eckles Dean, King Simon, Naaman Mor, Nair Rahul, Spasojevic Mirjana, Hui-I Yang Jeannie. 2006. ZoneTag: Designing Context-Aware Mobile Media Capture to Increase Participation. PICS '06. p. 3.
- [76] Naaman Mor, Nair Rahul. ZoneTag's Collaborative Tag Suggestions: What is This Person Doing in My Phone?. IEEE MultiMedia, 15 (2008) 3. pp. 34-40
- [77] Heinisuo Rami, Rauta Ilkka. PHP ja MySQL : tietokantapohjaiset verkkopalvelut. 4. painos. Helsinki 2007, Talentum. 412p.
- [78] Mobiililaitteita [WWW]. [viitattu 04.12.2009] saatavissa:
<http://www.uta.fi/~as63593/graksa/mobiililaitteita.htm>
- [79] The Python Wiki [WWW]. [viitattu 15.01.2009] saatavissa:
<http://wiki.python.org/moin/>
- [80] Scheible Jürgen, Tuulos Ville. Mobile Python – Rapid Prototyping of Applications on the Mobile Platform. England, 2007, John Wiley & sons, Ltd. 348p.
- [81] PHP time() Function [WWW]. [viitattu: 15.01.2010] saatavissa:
http://www.w3schools.com/php/func_date_time.asp
- [82] What is the Document Object Model? [WWW]. [viitattu: 15.01.2009] saatavissa: <http://www.w3.org/DOM/#what>
- [83] PHP utf8_decode() Function [WWW]. [viitattu: 15.01.2009] saatavissa:
http://www.w3schools.com/php/func_xml_utf8_decode.asp
- [84] Cherubini M., Anguera X., Oliver N., de Oliveira R. Text versus Speech: A Comparison of Tagging Input Modalities for Camera Phones. MobileHCI'09, Bonn, Germany September 15-18, 2009. Pp. 1-10.

- [85] Johdatus tilastolliseen hahmontunnistukseen (Oulun yliopisto) [PDF]. [viitattu: 16.01.2009] saatavissa:
<http://www.ee.oulu.fi/research/tklab/courses/521497S/pruju/Chapter1.pdf>
- [86] Kärkkäinen Henna. Hahmontunnistus kuvankäsittelyssä. Essee. Sipoo 2006. Teknillinen korkeakoulu, Informaatioverkostojen studio 4. 10p.

LIITE 1

Tagikysely:

Syötä alla oleviin kuviin mielestäsi parhaimmat tagiehdotukset.
Tagit ovat avainsanoja, jotka kuvaavat kuvan sisältöä.
Esimerkiksi Espanjan härkätäisteluja koskevassa kuvassa
tageja voisi olla "Espanja", "härkätäistelu", "härkä" ja "matador".



www.100apes.net

Tag1:

Tag2:

Tag3:

Tag4:

Tag5:



Tag1:

Tag2:

Tag3:

Tag4:

Tag5:



Tag1:

Tag2:

Tag3:

Tag4:

Tag5:

Lähetä

Kiitos osallistumisesta!